SF 298 MASTER COPY KEEP THIS COPY FOR REPRODUCTION PURPOSES

REPORT DOCUMENTATION PAGE

Form Approved
OMB NO. 0704-0188

AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	May 7, 2000	Final 01 Apr 96 - 31 Mar 00

4. TITLE AND SUBTITLE
Shape and Image Analysis using Neural Networks, Fractals and Wavelets.

5. FUNDING NUMBERS
DAAH04-96-1-0082

6. AUTHOR(S)
C.R. Rao
S.R.T. Kumara

7. PERFORMING ORGANIZATION NAMES(S) AND ADDRESS(ES)
The Pennsylvania State University
Center for Multivariate Analysis
417 Thomas Building
Dept. of Statistics
University Park, PA 16802

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Research Office
P.O. Box 12211
Research Triangle Park, NC 27709-2211

10. SPONSORING / MONITORING AGENCY REPORT NUMBER
ARO35518.63-MA

11. SUPPLEMENTARY NOTES
The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision, unless so designated by other documentation.

12 a. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for public release; distribution unlimited.

12 b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

A general theory and appropriate statistical analysis are developed for discrimination of objects by shape, i.e., by using features which are invariant to location, scale and orientation, in particular to reflection also. In the case of landmark data, features such as Euclidean distances between landmarks or angles of triangles after suitable triangulation are considered. When the objects do not have recognizable landmarks, as in the case of closed boundaries, we can use topological properties of points on the boundary at intervals of constant length as features. To deal with such cases, a new geometry of circular vectors with a suitably defined metric is developed. This enables the use of distance methods such as k-NN rule in pattern recognition. We have also concentrated on the extraction of features for representing shapes. As a generalization we have considered the signals from machining process and studied characterization using chaos and fractal analysis. We extended this work to represent shapes using wavelets, Fourier descriptors, fractal image compression and iterated functional systems. We have conducted a comparative analysis. In the contemporary internet world search engines need sophisticated techniques to search for images of interest based on shapes. We proposed a preliminary model and web bot based upon our shape analysis study to develop an image search engine for the www.

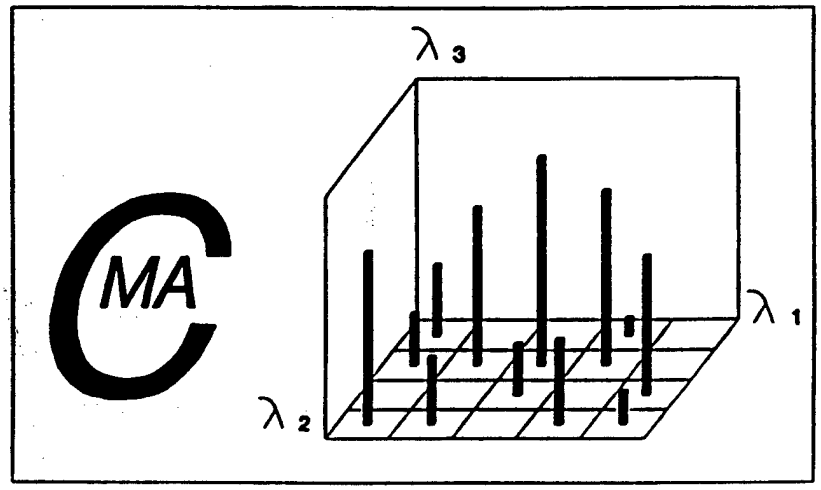
14. SUBJECT TERMS	15. NUMBER IF PAGES
Wavelet and Fourier descriptions, pattern recognition by shape Fractal dimension estimation, Multilayer perceptions, signal detection and estimation, Genetic algorithms, Iterative function systems.	16. PRICE CODE

17. SECURITY CLASSIFICATION REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UL

FINAL REPORT

1996-1999

RESEARCH WORK DONE UNDER SUPPORT
FROM ARMY RESEARCH OFFICE
GRANT No. DAAHO4-96-1-0082



CENTER FOR MULTIVARIATE ANALYSIS
STATISTICS DEPARTMENT
PENN STATE UNIVERSITY
UNIVERSITY PARK, PA 16802

PENNSTATE

Foreword

This is the final report of research activities at the Center for Multivariate Analysis, Statistics Department, Pennsylvania State University, under support from the Army Research Office, contract No. DA AH04-96-1-0082, during the period 1996-1999.

The results of research are reported in various Technical Reports issued by the Center for Multivariate Analysis and papers published in refereed journals and conference proceedings, lists of which are given in Appendices 3C and 3D of this report.

The research was undertaken by a team of senior scientists, post-doctoral research associates and graduate students. The principal investigators were C.R. Rao, Eberly Professor of Statistics, and S.R.T. Kumara, Professor of Industrial and Manufacturing Engineering. The senior collaborators were M.B. Rao, Professor of Statistics, North Dakota University, D.N. Shanbhag, Associate Professor, University of Sheffield, C.A.R. Murthy, Professor, Indian Statistical Institute, Calcutta, C. Balasubramaniam, Professor, Indian Institute of Technology, New Delhi, J. Pitman, Doctoral Candidate, Penn State University, Jaehong Suh, Doctoral Candidate, Penn State University and Shreesh Mysore, Doctoral Candidate, Penn State University.

The research was conducted on a broad spectrum of areas of applications of statistics in defense oriented problems. Some of the significant results are modified bootstrap methodology, signal detection and estimation, statistical analysis of shape for pattern recognition.

I wish to thank the ARO for the financial support which enabled the Center for Multivariate Analysis to put together a strong team of research staff to work on defense oriented problems.

Center for Multivariate Analysis
C.R. Rao
Penn State University
May 7, 2000

Table of Contents

1. Technical Objectives and Summary of Results 1

2. Significant Accomplishments 1

2.1 Shape Analysis 1

2.2 Signal Detection and Estimation 2

2.3 Application of Signal Analysis for Process Monitoring and Control 3

2.4 Feature Selection Using Neural Networks 4

2.5 Fractal Image Analysis 5

2.6 Web-based Image Analysis 6

2.7 Genetic Algorithms, Neural Networks, Fractals and linage Processing 6

2.8 Representation of Contaminated Chaotic Signals 8

2.9 Pattern Classification Using Annealing and Genetic Algorithms 9

2.10 Image Magnification and Compression Using Partitioned Iterated Function Systems 9

2.11 New Resampling Scheme for Bootstrap 10

3. Appendices

A. Research personnel of the Center for Multivariate Analysis (1996-1999) 11

B. Honors Received 13

C. List of papers published in refereed journals and conference Proceedings 14

D. List of technical reports with abstracts issued during the period 1996-1999 19

E. Participation in conferences 50

1. Statement of the Problems Studied

1.1 Technical Objectives and Motivation

Discriminant analysis based on multiple measurements is a fundamental tool in several practical applications ranging from disputed paternity and authorship of a manuscript to medical diagnosis character recognition, and detection and identification of signals. In many situations, the measurements made on an object depend on its location, scale and orientation. The technical problem involved is to extract features of objects which are invariant to the above transformations, which are called shape measurements and develop the statistical methodology for discrimination by shape.

Multifractal image analysis dealing with image representation and image compression is a problem of paramount importance in activities, and will assume tremendous importance in the future due to proliferation of new techniques such as remote and virtual collaboration in the internet. There is a need for a unified theory and framework for real-life implementation of new techniques like neural networks, wavelets, fractal analysis for image representation and subsequent classification.

In signal processing, high resolution signal parameter estimation is an important problem. In particular, the problems of detecting the number of signals and the estimation of direction of arrival of each signal are of special interest in military applications.

2. Significant Accomplishments

2.1 Shape Analysis

In a series of paper published in the Proceedings of the National Academy of Sciences (Rao and Suryavanshi (1996, 1998), Rao (1998, 2000)), appropriate statistical methods were developed for shape analysis.

Two types of objects were considered, one with identifiable landmarks and another characterized by a boundary without any landmarks. In the case of objects defined by landmarks, two types of measurements are considered. One is the set of Euclidean distances between all pairs of landmarks. Another is the set of all angles of triangles formed by all possible choices of three landmarks. In the case of distances data, statistical analysis is made on the differences of log distances which are invariant to SLRr (scale, location, rotation and reflection). Since the angles are invariant to SLRr, standard multivariate analysis could be applied.

It may be necessary to apply nonparametric or bootstrap methods if a suitable model cannot be found for the data.

In the paper Rao (2000), a new method of measuring the angles based on the orientation of a triangle is given which enables us to distinguish between objects if they differ in reflection. The method of angles appears to be most appropriate for shape analysis when objects are characterized by landmarks.

These methods are not applicable if there are no recognizable landmarks on an object. Such a situation arises when we want to discriminate between shapes of objects described by closed boundaries that do not have any natural landmarks. Some examples are contour coding, classification of chromosomes, interpretation of X-rays, scene analysis, identification of aircraft etc. A closed boundary of an object can be represented in many ways. We may start with some chosen point on the boundary and measure the coordinates or the curvature or any topological property of points around the boundary at intervals of constant boundary length. The vectors of such measurements on different objects are not comparable unless the starting point on each boundary is a well defined landmark. We can, however, represent an object by a closed string of measurements without specifying the starting value, which may be called circular vectors.

To analyze such measurements a geometry of circular vectors was developed by Rao (1998). The distance between two circular vectors and the sum of a number of circular vectors are suitably defined. These concepts enable the use of distance based methods in statistics such as k-NN rule and neural networks for pattern recognition and discrimination. Some examples are given to explain the computations.

Another approach is to construct a few functions of circular vectors that are independent of the starting point. Standard statistical methodology can be used on such invariant functions. Further work on the subject is in progress.

2.2 Signal Detection and Estimation

In signal processing, high resolution signal parameter estimation is an important problem. In particular, the estimation of the direction of the arrival of narrow band signals emitted by multiple sources has received much attention in signal processing literature. There are two problems associated with signal processing. One is the estimation of the number of signals and another is the estimation of the direction of arrival of each signal. The model is usually written in the form

$$X(t)=As(t)+n(t)$$

where X(t) is a p-vector of observations at p sensors, s(t) is a q-vector of signals and n(t) is noise n-vector, and A is a matrix whose elements are functions of directions of arrival of signals.

The problem of determining the number of signals (the parameter q in the above model) is solved by using model selection criteria developed at the Center for Multivariate Analysis (Bai and Rao, 1989 and Bai, Rao and Wu, 1999).

For the second problem of estimating the directions of arrival of signals, there are a number of algorithms such as ESPRT, TLS-SPRT, MUSIC and GEESE. Bai and Rao (1989) proposed a new method that gives more efficient estimates. Kundu (1998) made further improvements on this method. Simulation studies show that the proposed method yields estimators with much better performance than any known method in the signal processing literature.

2.3 Application of Signal Analysis to Manufacturing Process Monitoring and Control

The current research laid the foundation for robust techniques for signal analysis and feature extraction. We have applied these theoretical techniques to real time quality control of manufacturing processes.

In today's global manufacturing setting, on-line real-time control of manufacturing processes is deemed essential. On-line real-time quality control may be achieved through an integrated architecture - called *Intelligent Integrated Diagnosis* - consisting of monitoring, diagnosis and control schemes. The present day monitoring and control schemes do not adequately consider the complexities involved in a manufacturing process and are not based on a thorough understanding of the process mechanism and dynamics. Hence a new, systematic approach to developing monitoring and control schemes for manufacturing processes is very much in order.

A manufacturing process is a dynamic system whose inputs are determined by the specific actuations of process parameters such as speed and feed. Its outputs can be measured by the on-line sensors. Outputs commonly betray irregular, seemingly random, behavior. If this irregular behavior results from the characteristics of the process dynamics itself rather than the external noise, the process is said to exhibit chaos. In addition, if the process dynamics- can be adequately captured using a few state variables, then the dynamics is low-dimensional chaotic. Low-dimensional chaos is necessary for practical controllability of a manufacturing process.

The turning process on a lathe is a good example of a manufacturing process exhibiting low dimensional chaos. We have conclusively proved, through rigorous experimentation and a battery of hypothesis tests, that turning dynamics is low-dimensional chaotic.

This research concerns developing monitoring and control paradigms for low-dimensional chaotic processes based on thorough characterization of process dynamics. The paradigms are focussed towards addressing specific monitoring and control issues in the turning process.

The monitoring scheme is applied to continuous flank wear estimation. The monitoring scheme involves extracting feature's from sensor signals and relating them to the flank wear height using a recurrent neural network. Fractal dimensions extracted from the sensor signals serve as the signal features. Fractal dimensions serve as viable features because of the low-dimensional chaotic nature of turning dynamics.

The control scheme is applied to real-time chatter control. From the measured sensor signals a nonlinear model of turning dynamics has been developed. It accounts for various complex phenomena taking place in turning dynamics. The stochasticity of the dynamics and model uncertainties is characterized and modeled. Based on this model, a nonlinear control scheme based on combining differential geometric approach, chaos control methods and robust control methods has been developed. Rigorous testing of the control scheme through computer simulations reveal that the scheme is stable, robust to model uncertainties and on-line implementable. On-line implementation of this control scheme can be enforced through a neural network.

From a scholastic viewpoint, this research combines the previously disjoint areas of nonlinear and stochastic dynamics, and fractals on one hand with neural networks and signal processing methods such as wavelets on the other in order to develop practical monitoring and control schemes for manufacturing processes. This lays the foundation for nonlinear signal separation which is useful in shape analysis.

Details of the reported work appear in several papers listed in the technical reports (97-14) and in the paper in IJMT by Bukkapatnam et.al (1997). During the next phase of this research we will investigate the application of this methodology for monitoring real life physical systems. Currently we are using the Westland helicopter engine data to verify the methodology.

2.4 Feature Selection Using Neural Networks for Functional Approximation

The use of feed forward networks for function approximation has a history of theoretical results that can be traced back to the work of Kolmogorov. Practical applications for function approximation began to emerge only after the development of the back propagation training algorithm. There are several classical feature selection approaches suitable for small scale problems. However there is no unified theory of feature selection for function approximation using neural networks.

In this research, we systematically investigated the feature selection problem with a polynomial function. We varied the inputs and tried training the neural networks, with different topologies. We used raw data as well as rescaled data in our analysis. Our findings are summarized in the following:

1. When the data was rescaled, it was found that the performance seemed to improve with an increase in the order of the function. On the other hand, when the data was not rescaled, the network performance seemed to improve with an increase in the order of the function. This was explained by the fact that the values being estimated in this case were closer to zero.

2. When the data was rescaled, the use of a larger number of features improved the performance of the network in terms of the number of iterations required for convergence. When the data was not rescaled, the accuracy of the neural network was best when the number of input features were the fewest. However, the functions having very high orders, the number of input features used has no effect on the performance.

3. The performance of the neural network improved when the number of training samples was increased. The number of hidden nodes being used did have an effect on the performance. However, using 10 hidden nodes seemed to be a good choice.

4. To determine the optimal set of input features, the errors obtained in retrieval were summed over the different functions tested. In the case where the data was not rescaled, the optimal number of features was found to be 2, and these features were simply the 2 input variables. However, when the data was rescaled, no particular set of features could be selected as an optimal. In this case, the network performance seemed to improve as the number of input features increased.

5. From all the experiments performed, it was noticed that the mean square errors obtained were very large for those functions whose derivatives took high values. The functions having high orders contained regions where the slopes or derivatives were large and the value being estimated was close to 1. These regions were the most difficult to approximate. Therefore, it was decided to use an alternate method of measuring the neural network performance. A signal-to-noise ratio was used to evaluate the network performance. In this case, the signal was the value being estimated and the noise was the error. This ratio was plotted for different cases and it was found that the network was performing fairly well in all cases on this basis. It must be mentioned here that the signal-to-noise ratio results did not indicate an improvement in the network performance. This was merely an alternate method of evaluating the network performance.

The details of the work can be found in the technical report 97-19. We will use different functions to further test the ideas of function approximation. During the next year we will arrive at a theoretical means for feature selection.

2.5 Fractal Image Analysis

When an image is digitized and stored in the computer, there is always a loss of information with respect to the object it represents. Image analysis that encompasses image representation, image compression, image classification, and object recognition provides means to study an image in the context of the object it represents. The most fundamental issue in image analysis is as follows: How to develop a parsimonious representation of these images so that these can be used in conjunction with certain macro descriptors for classifying the objects using the images? Techniques to address this issue are grouped under the umbrella of image representation and classification.

In this research, we addressed image representation and classification. We considered black and white images without and with grey scaling. Typically boundary images such as hand profiles can be represented using a binary code. On the other hand, composite images such as a scene or a face could be represented by a particular gray scale.

All these images, over certain scales, exhibit *self-affinity* or *scale invariance,* i.e. the properties (usually the intensity distribution) over different scales or resolutions remains very similar. Our approach is to exploit this self-affine structure of the images to develop parsimonious image representations and thereby develop neural network-based classifiers. We conjecture that (1) the signature of the profiles of boundary images, and (ii) the pattern of gray scale images carry some self-similar structure through a range of macro-scales. We therefore propose fractal analysis to extract the scale-invariant properties of images.

For the boundary images, in the absence of any other attributes, only information available to us is the image which is mostly black/white. This signal (image) is topologically equivalent to a one-dimensional curve corresponding to the signal boundaries. The spatial variation of the curvature of this profiles, we speculate, follows a self-affine pattern similar to a *multifractal* entity.

On similar lines, for grey-scale images, the signal image is equivalent to a two-dimensional surface whose spatial representation follows a self-affine pattern similar to a *multifractal* entity.

We used wavelets-based multifractal representation and fractal image compression methods for image representation. Generalized fractal dimensions extracted from the parsimoniously represented images serve as signal features. Based on these features, we are currently developing neural network classifiers to perform both *group classification* and *individual classification.*

2.6 Web-based Image Analysis

We concentrate on the world wide web (or simply the web) as the platform for application of our fractal image analysis methodology. We illustrate the potential applicability of fractal image analysis using the following examples of arriving at a blueprint model of a house using the web. In order to extract all the available models of houses, a web-bot should be designed such that it recognizes house models from all the images it searches. In other words, the web-bot should be equipped with the capability to recognize an image to be a house model. The next step is to classify the recognized models into various categories in terms of certain elementary attributes like number of rooms, plinth shape, number of stories and roof-type.

Now, according to an individual's tastes and requirements, the closest matching model can be retrieved from among the retrieved models. If need be, certain algebraic operations may be performed on these models to exactly match the requirements.

The proposed scheme is shown in Figure 1. We state that the scheme in principle has a primitive predecessor in QBIC (Query By Image Content) where attributes such as color, texture and size are used as features for image classification as opposed to multifractal features in the proposed work.

A web-bot is released at regular periods of time to gather various images. Such image gathering may be accomplished through a contents-based search. The images thus gathered are filtered using a two-stage neural network scheme. In the first level, only the images that correspond to house models are retained and the rest are filtered-out. Next all the retained house model images are spatially organized in a self-organizing map so as. to enable fast retrieval based on values of certain attributes, and then stored in a compressed image format. This data base of images is labeled as house model repository in figure 1.

Depending on user requirements, a closest match is selected from among all house models in the repository. In addition, binary operations are performed on different images to provide the user with a satisfactory house model.

2.7 Genetic Algorithms, Neural Networks, Fractals and linage Processing

1 . *Genetic algorithms:* Genetic algorithms are stochastic search techniques based upon a natural genetic system. Three main processes of a natural genetic system are selection, reproduction and mutation. The underlying concepts of these processes are incorporated in genetic algorithms. Genetic algorithms are known to provide near optimal solutions to complex optimization problems. It has been shown recently that the elitist model of genetic algorithms will always

provide the optimal solution as the number of generations tend to infinity, whatever may be the initial population.

We have started using genetic algorithms for curve fitting. The genetic algorithms are applied to the problem of fitting the least squares piece-wise linear function where the knot locations are unknown (Tech. Rept. 97-15). The effectiveness of the procedure has been demonstrated successfully. This method will be extended to the case of finitely many piecewise linear functions where the number of such functions is unknown. Ultimately, we shall try to extend it to the general curve fitting problem.

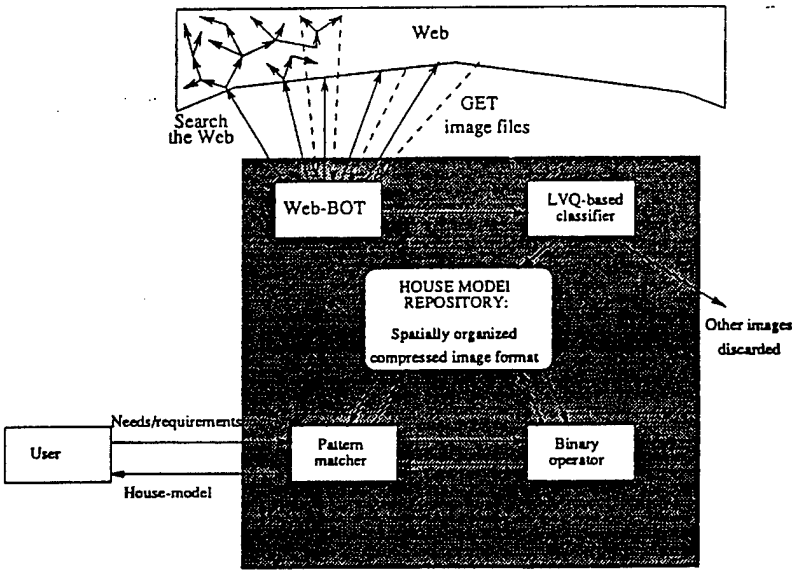


Figure 1: Illustration of the proposed application of fractal image analysis.

2. *Neural Networks:* An investigation has been made to find a connection between fractals and neural networks (Tech. Rept. 97-01). Gradient descent technique is the basis of several artificial neural network models, including the Multilayer Perceptron (MLP). Fractals are usually generated with the help of contractive maps. Contractive maps, when iterated repeatedly on any compact set, gives rise to a fractal under a few conditions.

It has been shown that the gradient descent technique is a contractive map, under some conditions. This established the relationship between neural networks and fractals in general. It is a matter of further research to utilize the above relationship to formulate better neural network models.

An investigation to classify pixels in satellite imagery has been conducted. A new LVQ model using exponential membership function has been suggested in this regard. This model (Tech. Rept. 96-09) has been successfully applied to artificial data sets and also the IRS (Indian remote sensing satellite) imagery with spatial resolution 36.25m x 36.25m.

3. Image Processing: Hough Transform is generally used in Image Processing for detecting line segments in images. We have used Hough transform for detecting regions (Tech. Rept. 96-15) in remotely sensed images. A region in an image is defined to be a union of line segments. The

detection of these line segments automatically gives rise to the detection of regions in a digital image. Thus, a new segmentation procedure for detecting regions in an image is developed. The procedure has been successfully demonstrated on an IRS image.

4. Fractals: A new procedure for fractal image compression using genetic algorithms has been developed. This procedure reduced the number of computations for finding the appropriate domain block and appropriate transformation for a range block in Jacquin's methodology with the help of genetic algorithms. The compression ratios and peak signal to noise ratio values are found to be comparable with the earlier methods.

The fractal coding scheme has also been used for image magnification. A 512 x 512 image has been constructed from a 256 x 256 image using the fractal code of the 256 x 256 image (Tech. Rept. 97-05). This image interpolation procedure is experimentally shown to compete with the other image interpolation procedures.

The fractal based image compression algorithm provides different compression ratios for different sizes of the range block. Different compression ratios provide different distorted images. We are going to study the relationship between the compression ratios and distortion measure values and try to model their relationship.

5. Fractals *and Genetic Algorithms:* Genetic algorithms are stochastic search techniques. The elitist model of genetic algorithms ensures that the optimal solution is reached as the number of iterations goes to infinity. For a given problem, whichever initial population that one starts with, the elitist model of genetic algorithms will always yield the optimal as the number of iterations goes to infinity.

Contractive maps and iterated function systems are used to construct fractals. Attractor is the limit of functions applied repeatedly on any non-empty compact set. Every iterated function system yields a unique attractor. For a given iterated function system, whichever initial set that one starts with, repeated application of the system will always yield the same attractor. Our aim is to find a relationship between these two methodologies. We have conducted preliminary studies to understand these relationships.

2.8 Representation of Contaminated Chaotic Signals

Signal representation is an essential step in many engineering signal processing applications such as pattern recognition and state estimation. The key task of signal representation is to find a basis to parsimoniously represent a given signal. In addition, the basis should (i) capture essential signal features such as discontinuities, (ii) match the smoothness of the signal, (iii) accommodate rapid fluctuations in the signal, (iv) cater to the stochasticity and the distribution of the signal, and (v) be usable on line.

Methods are available such as Karhunen-Loeve (KL) expansion, Fourier and wavelet coefficients to represent signals under restrictive conditions such as stationarity. None of these methods explicitly address the issue of optimal representation of chaotic signals. Realistically contaminated chaotic signals are more common than stationary signals.

Methods are developed in Bukkapatnam, Kumara and Lakhtakia (1998) to find sub optimal scaling functions using the local eigenfunctions to represent signals emanating from a chaotic process contaminated with low-level noise.

The methodology consists of (1) generating artificial ensembles from a pseudo-probability space constructed from the measured signal, (ii) extracting local eigenfunctions from the pseudoprobability space and (iii) using the extracted local eigenfunctions for suboptimal wavelet packet (WP) representation.

The application of this novel representation method to actual acoustic emission (AE) signals from a turning process (in machining) reveals the superiority of this method over the existing signal.

2.9 Pattern Classification Using Annealing and Genetic Algorithms

Methods for finding boundaries approximated by piece wise linear segments (hyperplanes) for classifying patterns in a p-dimensional vector space have been developed using Genetic and Simulated Annealing Algorithms (GA and SA). They involve generation and placement of a set of hyperplanes in the feature space to provide maximum separation of the clusters of patterns representing different classes or to minimize the frequency of misclassification when the hyperplanes are used as decision boundaries for discriminating between classes.

GA and SA are search algorithms for solving complicated optimization problems which have been found to outperform classical methods such as gradient descent search when the search space is large, complex and multimodal. SA has its foundation in statistical mechanics, which studies the behavior of a very large system of interacting components in thermal equilibrium. GA has its foundation in genetic concepts of chromosomes, crossover and mutation.

In two Technical Reports (Bandyopadhyay, Murthy and Pal (1998a, 1998b)) of the Center for Multivariate Analysis, SA and VGA (modification of GA using variable size of strings) have been used in pattern recognition problems for discriminating between vowel sounds. Both of them perform as good and often better than the standard methods of pattern recognition like Bayes maximum likelihood classification (BC), Multi Layer Perceptron (MLP), and k-NN rule.

2.10 Image Magnification and Compression Using Partitioned Iterated Function Systems

A new technique for image magnification using the theory of fractals is proposed. The technique is designed assuming self transformability property of images. In particular, the magnification task is performed using the fractal code of the image instead of the original one resulting in a reduction in memory requirement. To generate the fractal codes, Genetic Algorithm with elitist model is used which greatly decreases the search for finding self similarities in the given image. The article presents both theory and implementation of the proposed method. A simple distortion measure scheme and a similarity criterion based on just noticeable difference have also been proposed to judge the image quality of the magnified image. Comparison with one of the most popular magnification techniques, the nearest neighbor technique, is made.

The technique of image compression Iterative Function System (IFS) is known as fractal image compression. An extension of IFS theory is Partitioned or local Iterative Function System (PIFS) for coding the gray level images. Several techniques of PIFS based image compression have already been proposed by many researchers. The theory of PIFS appears to be different from the theory of IFS in the sense of application domain. The present article discusses some basic differences between IFS and PIFS and provides a separate mathematical formulation for the existence of attractor of partitioned IFS. In particular, it has been shown that the attractor exists and it is an approximation of the given target image. The experimental results have also been presented in support of the theory. The experimental results have been obtained by using a GA based PIFS technique developed in earlier papers.

2.11 New Resampling Scheme for Bootstrap

In the usual bootstrap, sampling of the observed data is done with replacement. In such a scheme, there is a possibility of very few distinct sampling units occurring in the bootstrap sample, which may cause some problems in the computation of bootstrap statistics. To avoid this a new sampling scheme is suggested where samples are drawn with replacement until a certain number of distinct units appear. The second order properties of such a sampling scheme is established under some general conditions.

This new bootstrap method is used in testing hypotheses based on shape measurements.

Appendix A: Research Personnel of the Center for Multivariate Analysis (1996-1999)

Faculty members and graduate students who participated in seminars, discussions and research work supported by the Army Research Office

From Statistics

1. C.R. Rao (Principal Investigator)
   Eberly Professor of Statistics and Director of the Center for Multivariate Analyais, Statistics Department, Pennsylvania State University.

2. M.B. Rao
   Professor of Statistics, North Dakota University, Fargo. Has a Ph.D. in Statistics and is the author of two books and 100 research papers.

3. D.N. Shanbhag
   Reader in Statistics, University of Sheffield, U.K. Has a Ph.D. in Statistics and is the author of one book and 90 research papers.

4. D.N. Kundu
   Professor of Statistics, Indian Institute of Technology, Khargpur, India. Has a Ph.D. in Statistics and is the author of 70 papers on signal detection, estimation and target tracking.

5. C. Balasubramaniam
   Professor, Indian Institute of Technology, New Delhi, India. Has a Ph.D. and is the author of 25 research publications.

6. C.A. Murthy
   Professor, Indian Statistical Institute, Calcutta. Has a Ph.D. and is the author of 30 papers on optimization problems, genetic algorithms and image compression.

7. Jennifer Pitman
   Graduate student supported by a grant from the Army Research Office. Was awarded a Ph.D. degree on the basis of a thesis on Adpative splines and genetic algorithms for optimal statistical modeling.

From Industrial Engineering

1. S.R.T. Kumara, Professor of Industrial and Manufacturing Engineering. Has more than 50 publications in process monitoring and diagnosis with emphasis on the application of advanced signal processing techniques.

2. Jaehong Suh, Doctoral Candidate, Industrial and Manufacturing Engineering. Is currently working on Condition Based Maintenance of heavy equipment with focus on wavelets as a representation mechanism.

3. Shreesh Mysore, Doctoral Candidate, Industrial and Manufacturing Engineering. Is currently working on Nonlinear dynamics with emphasis on chaos modeling. Completed MS thesis on Shape analysis.

4. Korinne Guza, Masters Candidate, Industrial and Manufacturing Engineering. Working on a comparative analysis of wavelets and Fourier descriptors for shape analysis.

Appendix B: Honors Received 1996-1999

C.R. Rao, Principal Investigator, received the following honors and awards during the period 1996-2000.

1996	D.Sc. (Hon. Causa), University of Guelph, Canada
1997	D.Math (Hon. Causa), University of Waterloo, Canada
1998	Doutor (Hon. Causa), University of Brasilia, Brazil
1999	D.Sc. (Hon. Causa), Athens University Economics and Business, Greece
2000	D.Sc. (Hon. Causa), Kent State University, Ohio, USA
1997	Foreign Member, Lithuanian Academy of Sciences, Lithuania
1998	Honorary Fellowship, International Indian Statistical Association
1996	Mahalanobis Birth Centenary Gold Metal; Awarded by the Indian Science Congress Association
1997	Distinguished Achievement Model of the ASA Section on Statistics and Environment
2000	Emmanual and Carol Panzen Prize for Statistical Innovation
2000	International Conference on Statistics: Reflections of the Past and Visions for the Future, in honor of C.R. Rao, on the occasion of his 80th birthday.
1997	Distinguished Alumni Award of the Delhi University
1997	Inducted into the Hall of Fame of the National Institution of Quality and Reliability, Chennai Branch, India
1998	Distinguished Ecology Award by the International Association for Ecology
1999 by the	National Award to young statisticians to be given once in 2 years was instituted Government of India in honor of C.R. Rao

Appendix C: List of Papers Published in Refereed Journals and Proceedings

C.1 C.R. Rao (Principal Investigator)

1996

a) (with D.N. Shanbhag). A note on a characteristic property based on order statistics. *Proceedings of the American Mathematical Society, Vol 124, 1, 299-302.*

b) Seven inequalities in statistical estimation theory. *Student, 1, 149-158.*

c) Empirical Bayes and hierarchical Bayes procedures in simultaneous estimation of parameters. In *Advances in Biometry* (Eds P. Armitage and H.A. David), Wiley, *115-130.*

d) (with L.C. Zhao). Law of the iterated logarithm for empirical cumulative quantile regression functions. *Statistica Sinica 6, 693-702.*

e) (with Shailaja Suryawanshi). Statistical analysis of shape of objects based on landmark data. *Proc. National Academy of Sciences 93, 12132-12136.*

f) Uncertainty, statistics and creation of new knowledge. *Chance, 9, 5-11.*

g) Principal component and factor analyses. *Handbook of Statistics Vol 14,* Eds. G. S. Maddala and C.R. Rao, North Holland, *489-506.*

1997

a) (with R. Mukerjee). Comparison of LR, Score and Wald statistics in a non iid setting. *J. Multivariate Analysis, 60, 99-110.*

b) (with L.C. Zhao). A limiting distribution theorem. In *Festschrift for Lucien Le Cam,Research papers in Probability and Statistics,* Eds D. Pollard, E. Torgerson and G.L. Yang, pp.325-336.

c) (with Z.D. Bai and Y. Wu). Robust inference in multivariate linear regression using difference of two convex functions as the discrepancy measure. *Handbook of Statistics 15,* Robust Inference, Eds. G.S. Maddala and C.R. Rao, North Holland, 1-19.

d) (with G.S. Maddala). Future directions of robust inference. *Handbook of Statistics,* Vol 15, Robust Inference, Eds. G.S. Maddala and C.R. Rao, North Holland, 661-675.

e) (with D.N. Shanbhag). Extensions of a characterization of an exponential distribution based on a censored ordered sample. In *Advances in the Theory and Practice of Statistics: A Volume in Honor of Samuel Kotz* (Eds. N.L. Johnson and N. Balakrishnan). John Wiley, 431-440.

f) (with N. Balakrishnan). A note on the best linear unbiased estimation based on order statistics. *American Statistician,* 51, 181-185.

g) Pre and post least squares. In *Proceedings of Statistical Challenges in Modern Astronomy* (Eds. G.J. Babu and E. Feigelson), Springer-Verlag, 3-13.

h) (with N. Balakrishnan). Large sample approximations to best linear unbiased estimation and best linear unbiased prediction based on progressively censored samples and some applications. In *Advances in Statistical Decision Theory and Applications.* In honor of Shanti Gupta (Eds. S. Panchapakesan and N. Balakrishnan), Birkhauuser, 431-444.

i) (with P.K. Pathak and V.I. Koltchinskii). Bootstrap by sequential sampling. *J. Statistical Planning and Inference,* 64, 257-281.

j) A cross disciplinary approach to teaching of statistics. *Proc. 51st Session of the Int. Stat. Institute,* Istanbul.

k) An alternative to correspondence analysis using Hellinger distance. *Proc. Int. Symp. on Contemporary Multivariate Analysis and its Applications,* Hong Kong, A11 -A29.

1998

a) (with V.K. Srivastava and H. Toutenburg). Pitman nearness comparisons of Stein-type estimators for regression coefficients in replicated experiments. *Statistical Papers,* 39, 61-74.

b) (with S. Suryawanshi). Statistical analysis of shape through triangulation of landmarks: A study of sexual dimorphism in hominids. *Proc. Nat. Academy of Sciences,* 95, 4121-4125.

c) (with N. Balakrishnan). Order statistics: An introduction. In *Handbook of Statistics,* Vol. 16, Order Statistics: Theory and Methods (Eds.: Balakrishnan, N. and Rao, C.R.), Elsevier, pp.324.

d) (with D.N. Shanbhag). Recent approaches to characterizations based on ordered statistics and record values. In *Handbook of Statistics,* Vol 16, Order statistics: Theory and Methods (Eds.: Balakrishnan, N. and Rao, C.R.), Elsevier, pp.231-256.

e) Geometry of circular vectors and pattern recognition of shape of a boundary. *Proc. Nat.Academy of Sciences,* 95, 12783-12786.

f) (with Hyder Ali). An overall test for multivariate normality. *Student,* 2, 317-324.

g) (with Z.J. Liu). M-estimation in a multivariate regression model under a convex-contoured discrepancy function. J. *of Combinatorics, Information and Systems Sciences,* 23, 1-9.

h) (with D.N. Shanbhag). Further versions of the convolution equation. *J. Ind. Soc.. Ag. Statist.* 51, 361-378.

i) (with D.N. Shanbhag). Recent approaches to characterizations based on order statistics. *Handbook of Statistics* 16, 231-256.

1999

a) (with Z.D. Bai and Y. Wu). Model selection with data oriented penalty. *J. Statistical Planning and Inference,* 77, 103-118.

b) Statistics: A technology for the millennium. *Internat. J. Math. & Statist. Sc.*

c) (with Z.D. Bai, Y. Wu, M. Zen and L. Zhao). The simultaneous estimation of the number of signals and frequencies of multiple sinusoids when some observations are missing: I. Asymptotics. *Proc. National Academy of Sciences,* 96, 11106-11110.

d) (with G.J. Babu and P.K. Pathak). Second order corrections of the sequential bootstrap. *Annals of Statistics*, 27, 1666-1683

Books Published During 1996-1999

C.R. Rao: Statistics and Truth (Second Enlarged Edition), World Scientific, Singapore.

C.R. Rao and M.B. Rao: Matrix Algebra and Its Applications to Statistics and Economics, World Scientific, Singapore.

Books Edited During 1996-1999

C.R. Rao and S. Ghosh: Handbook of Statistics, Vol. 13.

C.R. Rao and G.S. Maddala: Handbook of Statistics, Vol. 14.

C.R. Rao and G.S. Maddala: Handbook of Statistics, Vol. 15.

C.R. Rao and N. Balakrishnan: Handbook of Statistics, Vol. 16.

C.R. Rao and N. Balakrishnan: Handbook of Statistics, Vol. 17.

C.2 M.B. Rao (Senior Research Associate)

Papers Published (Refereed): 1996-1999

a) "On Improving the Quality of Minimal Incomplete Block Designs," *Journal of Applied Statistical Science, 4, 1996, 215-230* (with Manzoor Hussain and Theo Ogunyemi).

b) "A Note on the Iterated Logarithm for Weighted Sums of Independent Identically Distributed Random Variables," in *Convergence in Probability and Ergodic Theory,* Editors: V.Bergelson, P. March, and J. Rosenblatt, Walter de Gruyter, NewYork, *1996, 273-276* (with Deli Li).

c) "A Generalization of the Jerusalem Ticket Problem," *The Mathematical Scientist, 22, 1997, 37-44* (with V.V. Bapeswara Rao).

d) "Can One Load a Set of Dice so that the Sum is Uniformly Distributed?" *Mathematics Magazine, 70, 1997, 132-134* (with Warren Shreve and Guantao Chen).

e) "Result on a *2X2* Survival Experiment," *Mathematical Biosciences,* 146, 1997, 63-73 (with Ramin Bahreini-Arani).

f) "On S- and MS-optimality of Minimal Incomplete Block Designs," *Journal of Statistical Planning and Inference, 73, 1998, 189-195* (with Theo Ogunyemi).

g) "A Note on E-optimal Minimal Block Designs under Mixed Effects Model," *Journal of Applied Statistical Science, 6, 1998, 163-168* (with Theo Ogunyemi).

Papers Published (Proceedings): 1996-1999

a) "Interval-Censored Type 11 Plan," *Proceedings of the Section on Physical and Engineering Sciences,* American Statistical Association, *1996, 162-165* (with Surekha Mudivarthy and Rupa Mitra).

Papers Accepted for Publication

a) "The Law of the Iterated Logarithm and Central Limit Theorem for L-Statistics," *Journal of Multivariate Analysis* (with Deli Li and R.J. Tomkins).

b) "On Minimal S-Optimal Row-Column Designs," *Journal of Statistical Planning and inference* (with C.R. Rao and Theo Ogunyemi).

c) "Some Gambling Strategies: An Evaluation," *The Mathematical Scientist* (with Surekha Mudivarthy).

d) "On Improving the Quality of Row-Column Designs," *Journal of Applied Statistical Science* (with Theo Ogunyemi).

Papers Submitted for Publication

a) "On Weak-Consistency in Linear Models with Equi-Correlated Errors," *Statistics* (with Xinwei Jia).

b) "A Note on Inter-Class Correlation: A Case for Unbalanced Regression," *The Annals of Institute of Statistical Mathematics* (with Marie Palta and Raja Velu).

C.3 S.R.T. Kumara (Faculty Associate)

Journal Publications

1. Bukkapatnam, S., A. Lakhtakia, and S. Kumara. 1996.Conceptualization of Chaos Theory Based Optimal Cutting Tool Chatter Control. *Speculations in Science and Technology* 19: 137-148.

2. Bukkapatnam, S., A. Lakhtakia, and S. Kumara. 1997. Chaotic Neurons for On-line Quality Control in Manufacturing. *International Journal of Advanced Manufacturing Technology.* 13: 95-100.

3. Kim, T., and S. Kumara. 1997. Boundary Defect Recognition Using Neural Networks. *International Journal of Production Research* 35(9): 2397-2412.

4. Bukkapatnam, S., Kumara, S.R.T., and Lakhtakia, A. 1999. Analysis of Acoustic Emission Signals in Machining. *ASME Transactions in Manufacturing Science and Engineering.* 121: 568-576.

5. Bukkapatnam, S., S. Kumara and A.L. Lakhtakia. 2000. Tool Wear Estimation in Turning using Non-Linear Dynamics. *ASME Transactions on Dynamics and Control.* (Accepted - to appear).

6. Bukkapatnam, S., S. Kumara and A.L. Lakhtakia. 1999. Local Eigenfunction based suboptimal wavelet packet representation of contaminated chaotic signals. *IMA Journal of Applied Mathematics.* (Accepted - to appear).

7. Suh, J., Kumara, S.R.T., Mysore, S., Lang, D., and Garga, A. 1999. Wavelets Assisted Condition Monitoring of Heavy Equipment. *Annals of the International Institution for Production Engineering Research (Annals of CIRP).* (Accepted - to appear).

8. Kamarthi, S., S. Kumara, and P. Cohen. 2000. A Review of Important Tool Estimation Methods in Turning. *IIE Transactions.* (Accepted - to appear).

9. Kamarthi, S., S. Kumara, and P. Cohen. 2000. Application of Wavelets to AE Signal Processing in Metal Cutting. *ASME Transactions in Manufacturing Science and Engineering.* (Accepted - to appear).

Conference Presentations:

a) Advanced Signal Processing Techniques in Monitoring, STC-Machines Meeting, 49th CIRP General Assembly, Montreaux, Switzerland, August 1999. (Presenter)

b) Wavelets Assisted Condition Monitoring of Heavy Equipment, STC-Machines Technical Session, 49th CIRP General Assembly, Montreaux, Switzerland, August 1999.(Presenter)

TECHNICAL REPORTS: 1996

Issued by the Center for Multivariate Analysis

Supported by ARO Grant DAAH04-96-1-0082

96-01 C.R. Rao, P.K. Pathak and V.I. Koltchinskii. Bootstrap by Sequential Resampling.

Abstract: This paper examines methods of resampling for bootstrap from a survey sampling point of view. Given an observed sample of size n, resampling for bootstrap involves n repeated trials of simple random sampling with replacement. From the point of view of information content, it is well known that simple random sampling with replacement does not result in samples that are equally informative (see Pathak (1964)). This is due to different numbers of distinct observations occurring in different bootstrap samples. We propose an alternative scheme of sampling sequentially (with replacement each time) until k distinct original observations appear. In such a scheme, the bootstrap sample size becomes random as it varies from sample to sample, but each sample has exactly the same number of distinct observations. We show that the choice of $k=(1-e^{-1})n \approx .632n$ has some advantage, stemming from the observation made by Efron (1983) that the usual bootstrap samples are supported on approximately $.632n$ of the original data points. Using recent results on empirical processes, we show that main empirical characteristics of the sequential resampling bootstrap are asymptotically within the distance of order $n^{-3/4}$ of the corresponding characteristics of the usual bootstrap.

Accepted for *J. Statist. Plann. Inference.*

96-02 C.R. Rao. Uncertainty, Statistics and Creation of New Knowledge.

Abstract: How do we take decisions under uncertainty? How do we generalize from particular observed data to discover new phenomena or to advance natural knowledge? Is the process involved an art, technology or science? The paper discusses the methods of reasoning by which new knowledge is created and the associated philosophical issues and controversies. The role of statistics, the newly developed discipline of the present century, in this process is described.

Published in *Chance, Vol 9, 5-11 (1996).*

96-03 C.R. Rao and D.N. Shanbhag. Extensions of a Characterization of an Exponential Distribution Based on a Censored Ordered Sample.

Abstract: Dufour gave a conjecture on a characterization of an arbitrary exponential distribution based on a right-censored ordered sample. The conjecture was shown to be true by Leslie and van Eeden (1993) in the case when the number of censored observations is no larger than $(1/3)n-1$, and by Xu and Yang (1995) and Rao and Shanbhag (1995) when the number is less than or equal to $\max\{0,n-5\}$, where $n(\ 3)$ is the sample size. In the present paper we give extended versions of the results concerning the Dufour conjecture.

Submitted: *Kotz Volume.*

96-04 C.R. Rao. Principal Component and Factor Analyses.

Abstract: Principal component (PCA) and factor analyses (FA) are exploratory multivariate techniques used in studying the covariance (or correlation) structure of measurements made on individuals. The methods have been used by applied research workers in a variety of ways, from reducing high dimensional data to few functions of variables carrying the maximum possible information, grouping of similar measurements and detecting multicollinearity, to graphical representation of high dimensional data in lower dimensional spaces to visually examine the scatter of the data and detection of outliers.

The computations involved in these methods and the interpretation of results in different situations are discussed. The difference between PCA and FA, and the need to choose the appropriate technique in the analysis of given data are stressed.

It is shown that there is a close similarity between the growth curve models used in biometric studies and the arbitrage pricing theory model recently introduced in financial statistics.

Published in *Handbook of Statistics Vol 14, 489-506 (1996).*

96-05 S. Bukkapatnam, S.R.T. Kumara, A. Lakhtakia. Fractal Estimation of Flank Wear in Turning. Part-1: Theoretical Foundations and Methodology.

Abstract: In this two-part paper, a novel scheme of sensor-based on-line cutting tool flank wear estimation, called fractal estimation is developed, implemented and evaluated. This paradigm is unique in the sense that we extract fractal properties of sensor signals. The metric invariants of the sender signals called *fractal dimensions* are related to the instantaneous flank wear using a recurrent neural network to implement a *fractal estimator*. The

performance of the fractal estimator, evaluated using actual experimental data, establishes this scheme as a viable flank wear estimation paradigm. This methodology is general enough to be applied to many classes of estimation problems related to several manufacturing processes.

We have developed the necessary theoretical formalisms and obtained implementation experiences through the research on tool wear monitoring in turning. The feature extraction methods used in this work are vital to the image analysis research and form the foundation for our future work.

In the first part, theoretical foundations leading to the development of the fractal estimator are presented. New schemes of wavelet transform-based signal separation and fractal dimensions based feature extraction are described in detail.

Submitted: *ASME Transactions on Dynamics, Systems and Control.*

96-06 S. Bukkapatnam, S.R.T. Kumara, A. Lakhtakia. Fractal Estimation of Flank Wear in Turning. Part-2: Implementation Details.

Abstract: In Part-1 of this paper we described the theoretical foundations of fractal estimator development. The implementation aspects of the fractal estimator for on-line tool wear estimation in the turning process is described in this part. The results show that this paradigm can be successfully used for real-time quality control in manufacturing.

Submitted: *ASME Transactions on Dynamics, Systems and Control.*

96-07 V.I. Koltchinskii and Lang Li. A Bootstrap Test for Spherical Symmetry of a Multivariate Distribution.

Abstract: We suggest a test for spherical symmetry of a distribution in $\mathbf{R}^d$ with unknown center. The test statistic is based on the multivariate extensions of the distribution and quantile functions, recently introduced by Dudley and Koltchinskii (1992) and Chaudhuri (1996). We study the asymptotic behaviour of the sequence of test statistics for large samples and justify a version of bootstrap approximation of their distributions.

Submitted: *Statistica Sinica.*

96-08 C.R. Rao and D.N. Shanbhag. Further Versions of the Convolution Equation.

Abstract: Many authors seem to be unaware of the importance of the existence of the earlier papers of Choquet and Deny (1960) and Deny (1961) on the convolution equation, even when these have links with their own papers. Lau and Rao (1982) and Davies and Shanbhag (1987) among others

have established variants or extended versions of the results of Choquet and Deny (1960) and Deny (1961) and have given various applications of these. The recent monograph of Rao and Shanbhag (1994) should provide the reader with the relevant details of the literature in this connection. In the present paper, we make some further observations on the literature on integral equations, pointing out, in particular, that certain results of Laczkovish (1986) and Baker (1994) are essentially simple corollaries to or variants of the general theorem of Deny (1961) or the theorem of Choquet and Deny (1960).

Submitted: *London Mathematical Society.*

96-09 S. Bose and C.A. Murthy. A New LVQ Model.

Abstract: A new LVQ model has been proposed here. An exponential membership function has been considered in this regard. The performance of the new model in relation to other existing models has been studied experimentally with the help of an artificial data set as well as IRIS data. Finally the proposed algorithm is applied on a satellite image data. The proposed model has been found to provide satisfactory results with all these data sets.

Submitted: *Pattern recognition.*

96-10 C.R. Rao and Shailaja Suryawanshi. Statistical analysis of Shape of Objects Based on Landmark Data.

Abstract: Two objects with homologous landmarks are said to be of the same shape if the configurations of landmarks of one object can be exactly matched with that of the other by translation, rotation/reflection, and scaling. The observations on an object are coordinates of its landmarks with reference to a set of orthogonal coordinate axes in an appropriate dimensional space. The origin, choice of units and orientation of the coordinate axes with respect to an object may be different from object to object. In such a case, how do we quantify the shape of an object, find the mean and variation of shape in a population of objects, compare the mean shapes in two or more different populations, and discriminate between objects belonging to two or more different shape distributions. We develop some methods which are invariant to translation, rotation and scaling of the observations on each object and thereby provide generalizations of multivariate methods for shape analysis.

Published in *Proceedings of the National Academy of Sciences, vol 93, 12132-12136 (1996).*

96-11 C.R. Rao. Pre and Post Least Squares: The Emergence of Robust Estimation.

Abstract: The paper traces the history of unknown parameters when measurements are subject to error from the time of Ptolemy to Gauss and Laplace, the inventors of the method of least squares estimation (LSE). The modern theory of LSE started with the papers by Markoff and Aitken and later contributions by Bose and the author. The LSE's have some nice properties. However, they are found to be sensitive to outliers and contamination in the data. To overcome this defect, robust methods are introduced using measures of discrepancy between a measurement and its expected value which have a slower rate of growth than the squared value. A unified theory of robust estimation is described using the difference of two convex functions as the measure of discrepancy.

Submitted: *Proceedings of Statistical Challenges in Modern Astronomy.*

96-12 C.R. Rao. Canonical Coordinates for Graphical Representation of Multivariate Data.

Abstract: A general theory is developed for representing population profiles characterized by multiple measurements in a low dimensional Euclidean space. The basic inputs of the problem are a matrix of distances between population profiles and the weights to be attached to different population profiles. The derived coordinates in the reduced space are called canonical coordinates. A well known method for representing row or column profiles in a contingency table using a chisquare type distance between profiles is the correspondence analysis. It is suggested that a similar analysis based on Hellinger distance between profiles has some advantages and is better suited for studying the configuration of profiles.

An asymmetric biplot technique which is useful in interpreting differences in row (column) profiles in terms of column (row) categories is developed.

Submitted: *Saleh's Volume.*

96-13 C. Mudivarthy, M.B. Rao and R. Mitra. Interval-Censored Type II Plan.

Abstract: The Type II plan is used on many occasions in order to estimate the lifetime distribution of a product under investigation. Begin the process by choosing two positive integers r n. Select a sample of n units of the product, set them to work, and observe the units continuously until r units fail. The objective is to estimate the lifetime distribution of the product using the data on r failure times.

We want to offer a modification of this plan in response to a past

consultation problem. This problem arose from two diverse fields: one from engineering and the other from ornithology. In order to expedite observation of the r failure times, only periodic inspections were made. In such a plan, the exact failure of the r units will be unknown, but we will know how many units failed between each of the consecutive inspection times.

Submitted: *Proc. Amer. Stat. Association, Section on Engineering.*

96-14 M. Ziejewski, B. Anderson, M.B. Rao and M. Hussain. Energy Absorption for Short Duration Impacts.

Abstract: The objective of this study is to relate energy absorption characteristics to selected material properties and to establish a methodology that allows one to determine some of the material properties for maximum energy absorption. The finite element program DYNA-3D and its associated pre and post processors were used. The model used is a hollow square column. Five properties of the materials were included in the analysis: (i) Density (ii) Elastic Modulus (iii) Tangent Modulus (iv) Yield Strength, and (v) Poisson Ratio. The Response Surface Method in conjunction with the canonical analysis were employed to locate the optimum or near optimum levels of the properties and then to determine the equation of the response surface in an area near the vector of optimum levels. For the given levels of three out of five material properties used in the study, one can calculate the remaining two material property levels to achieve the near-optimal energy absorption.

Submitted: *Technical Paper Series No.961851 of the Society for Automotive Engineers, 1996.*

96-15 B. Uma Shankar, C.A. Murthy and S.K. Pal. A New Gray Hough Transform for Region Extraction from IRS Images.

Abstract:A technique using Hough transform is described for detection of homogeneous line segments directly from (i.e., without binarization of) gray level images. A definition of "region" in terms of these line segments, with constraints on its length and variance, is provided. The algorithm is able to extract gray regions irrespective of their shape and size. The effectiveness of the method is demonstrated on Indian Remote-sensing Satellite (IRS) images.

Submitted: *Pattern Recognition Letters.*

96-16 Ahmed A. Anu-Taleb, M. Bhaskara Tao. Interval-Censored Data.

Abstract: In estimating the lifetime distribution of a product, the following

periodic inspection plan is put forward. Select a random sample of n units of the product and set them to work. Choose an integer m 1, numbers $0<t_1<t_2<...<t_m$, and inspect the units at times t_1, t_2, ..., t_m. The data consist of number of units that fail between every two consecutive inspection times. The choice of m and numbers t_1, t_2, ..., t_m are discussed in this paper. Some graphical methods are presented to identify the parametric model of the underlying lifetime distribution based on the interval-censored data gathered. A new method of estimating the parameters of the lifetime distribution based on Linear Model Theory is presented and compared with the method of maximum likelihood.

Submitted: *Biometrika*.

LIST OF TECHNICAL REPORTS ISSUED BY THE CENTER FOR MULTIVARIATE ANALYSIS UNDER THE GRANT DAAH04-96-1-0082

Jan 1 - Dec 31, 1997

97-01 C.A. Murthy and Jennifer Pittman. Multilayer Perceptrons and fractals.

In this article, the mathematical relationship between the gradient descent technique and contractive maps is studied. A theoretical result regarding the eigenvalues of the Hessian matrix for the objective function is derived using contractive maps. Since the gradient descent technique is used in multilayer perceptrons, and contractive maps give rise to fractals, a theoretical relationship is thus established between multilayer perceptrons and fractals.

Submitted: *J. Information Science*

97-02 G.R.M. Borzadaran and D.N. Shanbhag. General characterization theorems via the mean absolute deviation.

The Chernoff-Borovkov-Utev inequality resulted due to earlier inequalities established by Chernoff (1981) and Borovkov and Utev (1983) respectively, giving bounds for the variance of functions of normal r.v.'s and leading to characterizations of normality. Subsequently, several analytic properties of variance bounds and other relevant results were established by others. Defining the median absolute deviation as E|X-med(X)| where med(X) is a median of the distribution of the random variable X, Freimer & Mudholkar (1991) gave a bound for the median absolute deviation of a certain real-valued function of an absolutely continuous random variable (w.r.t. Lebesgue measure) and Korwar (1991) presented an analogue of this in the discrete case; these authors, also, characterized the Laplace and a mixture of two Waring distributions via the respective bounds.
We extend these latter characterization theorems to the case where the distributions are not necessarily purely discrete or absolutely continuous, via the approach of Alharbi & Shanbhag (1996). The results in Freimer & Mudholkar (1991) and Korwar (1991) are now corollaries to this general result. Also, following Alharbi & Shanbhag (1996), we relate these results to Cox's representation for a survival function in terms of the hazard measure, as presented by Kotz & Shanbhag (1980). (The original version of the representation mentioned had appeared in Cox (1972).)

Submitted: *J. of Statistical Planning & Inference.*

97-03 A. Mitra and D. Kundu. Consistent method for estimating sinusoidal frequencies: a non-iterative approach.

97-04 D. Kundu and N. Kannan. Constrained maximum likelihood estimators for superimposed exponential signals.

Recently Kundu (1993a) has proposed a non-linear eigenvalue method for finding the maximum likelihood estimators (MLE) of the parameters of undamped exponential signals. It is known to perform better than the previously existing methods like FBLP of Tufts and Kumaresan (1982) or Pisarenko's method (Pisarenko; 1972), in the sense of lower mean squared errors. The solution in general depends on the roots of a polynomial equation. It is observed that the coefficients of the polynomial exhibit a certain symmetry. Since it is known (Crowder; 1984) that the MLE with constraints is more efficient than the unconstrained MLE, modified maximum likelihood method has been suggested to estimate the parameters under these symmetric constraints. It is observed in a simulation study that the mean squared errors of the constrained MLE are closer to the Cramer-Rao lower bound than the ordinary MLE in almost all situations.

97-05 C.R. Rao. A cross disciplinary approach to teaching of statistics.

Historically, the phenomenal growth of statistics as a separate discipline of great ubiquity was motivated by the need to solve practical problems arising in social, biological and natural sciences. Statistical methodology as we practice today involves acquisition of data, extraction of available information and taking optimal decisions under uncertainty. That the role of statistics as the logic and science of solving problems in other disciplines and that continued advances in statistics depend strongly on research stimulated by and directed at problems in other disciplines are not fully reflected in the educational programs for statisticians at the universities. It is suggested that more emphasis should be given to the interface between statistics and other disciplines through data-centered training in statistics.

Submitted: *Proc. of 51st Session of International Statistical Institute.*

97-06 C.R. Rao. An alternative to correspondence analysis using Hellinger distance.

In this paper, a general theory of canonical coordinates is developed for reduction of dimensionality in multivariate data, assessing the loss of information and plotting higher dimensional data in two or three dimensions for visual displays. The theory was applied to data in two way tables with variables in one category and samples (individual or populations) in the other. The method is applicable to data with continuous measurements on the

variables as well as to frequencies of attributes. An alternative to the usual correspondence analysis of contingency tables based on Hellinger rather than the chisquare distance is suggested. The new method has some attractive features and does not suffer from some inherent drawbacks resulting from the use of the chisquare distance and variable sample sizes for the populations in the correspondence analysis. The technique of biplots where the populations and the variables are represented on the same chart is discussed.

Submitted: *Contemporary Multivariate Analysis and Its Applications.*

97-07 Z.D. Bai, C.R. Rao and Y. Wu. Model selection with data-oriented penalty.

We consider the model selection or variables selection in the classical regression problem. In the literature, there are two types of criteria for model selection, one based on prediction error (FPE) and another on general information theoretic criterion (GIC). Each of these criteria uses a certain penalty function which is the product of the number of variables j in a submodel and a function C_n depending on n and satisfying some conditions to guarantee consistency in model selection. One of the important problems in such a procedure is the actual choice of C_n in a given situation. In this paper, we show that a particular choice of C_n based on observed data, which makes it random, preserves the consistency property and shows improved performance over a fixed choice of C_n.

Submitted: *J. Statistical Planning and Inference.*

97-08 N. Balakrishnan and C.R. Rao. Large-sample approximations to the best linear unbiased estimation and best linear unbiased prediction based on progressively censored samples and some applications.

In this paper, we consider the situation where a life-testing experiment yields a Type-II progressively censored sample. We then develop large-sample approximations to the best linear estimators for the scale-parameter as well as for the location-scale parameter families of distributions. Large-sample expressions are also derived for the variance and covariance of these estimators. There results are used further to develop large-sample approximations to the best linear unbiased predictors of future failures. Finally, we present two examples in order to illustrate the methods of inference developed in this paper.

Submitted: *The Volume in Honor of Shanti S. Gupta.*

97-09 S. Mitra, C.A. Murthy and M. K. Kundu. A Fractal based technique for

image magnification.

A new method for image magnification using fractal based technique is proposed. We call this new technique fractal image magnification. The technique is designed assuming the self transformability property of images. The technique described here also utilizes Genetic Algorithm with elitist model that greatly decreases the search for finding the selfsimilarities in the given image. The article presents theory and implementation of the proposed method. A simple distortion measure scheme is also proposed to judge the image quality of the magnified image. Comparison with one of the most popular magnification techniques, the nearest neighbor technique, is made.

Submitted: *Handbook of Statistics, Vol 19.*

97-10 D. Kundu. Asymptotic theory of the least squares estimators of sinusoidal signal.

The consistency and the asymptotic normality of the least squares estimators are derived of the sinusoidal model under the assumption of stationary random error. It is observed that the model does not satisfy the standard sufficient conditions of Jennrich (1969), Wu (1981) or Kundu (1991). Recently the consistency and the asymptotic normality are derived for the sinusoidal signal under the assumption of normal error (Kundu; 1993) and under the assumptions of independent and identically distributed random variables in Kundu and Mitra (1996). This paper will generalize them. Hannan (1971) also considered a similar kind of model and established the result after making the Fourier transform of the data for one parameter model. We establish the result without making the Fourier transform of the data. We give an explicit expression of the asymptotic distribution of the multiparameter case, which is not available in the literature. Our approach is different from Hannan's approach. We do some simulations to examine the small sample properties of the two types of estimators.

Submitted: *Statistics.*

97-11 G.R. Mohtashami Borzadaran and D.N. Shanbhag. Further results based on Chernoff-type inequalities.

In this paper, we address questions dealing with characterizations based on Chernoff-type moment inequalities and their variants and establish, via the approach of Alharbi & Shanbhag (1996), a general theorem extending, among others, various results of Cacoullos & Papathanasiou (1995a, 1995b).

Submitted: *Stat. and Prob. Letters.*

97-12 Ramin B. Arami and M.B. Rao. A result on a 2x2 survival experiment.

Lifetime data classified according to categorical variables under the proportionality of the hazard functions of response variables for various treatment combinations is assumed. The proposed model is a combination of Cox's proportional hazards model and ANOVA model. The existence of a solution to the marginal likelihood function is examined for the case of 2x2 two-way classification. We provide an easily verifiable condition for the existence of a unique estimate.

Submitted: *Mathematical Biosciences.*

97-13 Deli Li, M.B. Rao and R.J. Tomkins. The law of the iterated logarithm and central limit theorem for L-statistics.

The main idea in this paper is that we devise an effective way of combining the Smirnov's law of the iterated logarithm for empirical processes, and some well-known results of limit behavior of L-statistics to establish new results on the central limit theorem, law of the iterated logarithm, and strong law of large numbers, for L-statistics. We show further that this approach can be pursued profitably to obtain necessary and sufficient conditions for either almost sure convergence or convergence in distribution of some well-known L-statistics and U-statistics. A law of the logarithm for weighted sums of order statistics is stated without proof.

Submitted: *Annals of Statistics.*

97-14 S. Bukkapatnam, S.R.T. Kumara and A. Lakhtakia. Local eigenfunctions based suboptimal wavelet packet representation of contaminated chaotic signals.

We report a suboptimal wavelet packet (WP) representation of signals emanating from a chaotic attractor contaminated by low levels of noise. Our method - geared towards choosing a suboptimal scaling function to parsimoniously represent the signal - involves extracting local eigenfunctions using artificial ensembles generated from a *pseudo-probability space*, and using the extracted local eigenfunctions to develop a suboptimal scaling function. The application of our novel representation method to actual acoustic emission (AE) signals from the turning process reveals the superiority of these methods over the existing signal representations.

Submitted: *IMA Journal of Applied Mathematics.*

97-15 Jennifer Pittman and C.A. Murthy. Optimal line fitting using genetic

algorithms.

Genetic algorithms are computational techniques which, given an optimization problem, use elements of directed and stochastic search to find the "best" solution from the space of potential solutions. We apply GA's to the problem of fitting the minimum least-squares piecewise linear function to a set of data points in R^2, We assume that the number of pieces is known but the knot locations are unknown. The effectiveness of our algorithm is demonstrated with two examples. Results are found to be quite promising and encourage further research.

Submitted: *Pattern Recognition.*

97-16 Z.J. Liu and C.R. Rao. M-Estimation in Multivariate Regression Model Under a Convex-Contoured Discrepancy Function.

The M-estimation in a multivariate regression model is studied in this paper using a convex contoured discrepancy function (C-CDF). Most of the discrepancy functions considered by various authors can be expressed as C-CDF's. A general asymptotic theory of M-estimation is developed that includes strong consistency, asymptotic representation and asymptotic normality. Results of this paper generalize most of the results in the literature on M-estimation under minimal assumptions.

Submitted: *J. Statistical Planning and Inference (Volume in honor of J.N. Srivastava).*

97-17 S. Bukkapatnam, S.R.T. Kumara and A. Lakhtakia. Analysis of Acoustic Emission Signals in Machining.

Acoustic emission (AE) signals are emerging as promising means for monitoring machining processes, but the current understanding of their generation is very limited and existing techniques to analyze them are inadequate. In this paper, we present a novel methodology based on chaos theory, wavelets and neural networks, for analyzing AE signals. Our methodology involves a thorough signal characterization, followed by signal representation using wavelet packets, and state estimation using multilayer neural networks. Our methodology yields better performance than the existing methods of AE signal analysis.

Submitted: *International Journal of Production Research.*

97-18 S. Bukkapatnam, S.R.T. Kumara and A. Lakhtakia. The Neighborhood Method and its Coupling with the Wavelet Method for Signal Separation of

Chaotic TSD.

We report the development of two novel signal separation methods - (I) the neighborhood method (NM) and (ii) a modified wavelet method (MWM) - that seem to be adequate for chaotic time-series data (TSD) with small, uniform Lyapunov exponents. For the NM, a variant of shadowing methods used for signal separation of chaotic TSD, we establish theoretical bounds on performance under various noisy conditions and analyze its algorithmic complexity. Our MWM is an adaptation of Donoho's wavelet method (WM) to nonlinear, and possibly chaotic, TSD with multiplicative noise. It incorporates certain features of the NM thus outperforms both the NM as well as the WM. The MWM has lower algorithmic complexity that the NM, and is, therefore, more suitable for on-line implementation. Both methods were tested on chaotic TSD corresponding to the reconstructed Rossler attractor. A discussion on the applicability of both methods to TSD obtained from actual sensor signals is provided.

Submitted: *IEEE Transaction on Signal Processing.*

97-19 V. Ramachandran and S.R.T. Kumara. Feature Selection in Neural Networks for Function Approximation.

Function approximation is one of the major applications of neural networks and has been used in such areas as systems control and prediction in financial analysis. Despite the fact that extensive research has been conducted in feature selection, no clear guidelines have been established to select the minimal set of features to arrive at a function. In this paper, we investigate feature selection methods for function approximation using neural networks. Extensive experimentation is performed using a three-layer feed forward network to estimate a known set of function: polynomials of different orders. The number of input features to the networks and orders of polynomials are varied to study the behavior of the network. The errors in the estimation of the function are calculated in each case and are used to decide which set of features is the best for the given function approximation.

Submitted: *INFORMS Journal of Computing.*

97-20 C. Radhakrishna Rao and Shailaja Suryawanshi. Statistical Analysis of Shape Through Triangulation of Landmarks: A Study of Sexual Dimorphism in Hominids.

Two objects with homologous landmarks are said to be of the same shape if the configuration of landmarks of one object can be exactly matched with that of the other by translation, rotation/reflection, and scaling. In an earlier

paper, the authors proposed statistical analysis of shape by considering logarithmic differences of all possible Euclidean distances between landmarks. Tests of significance for differences in the shape of objects and methods of discrimination between populations were developed using such data. In the present paper, the corresponding statistical methodology is developed by triangulation of the landmarks and considering the angles as natural measurements of shape. This method is applied to the study of sexual dimorphism in hominids.

Submitted: *Proceedings of the National Academy of Science.*

97-21 C.R. Rao and Hydar Ali. An Overall Test for Multivariate Normality.

There are a number of methods in statistical literature for testing whether observed data come from a multivariate normal distribution with an unknown mean vector and covariance matrix. Let $x_1,...,x_n$ be a p-variate sample with sample mean and variance-covariance matrix *S*. Most of the tests are based on the result that $y_i = S^{-1/2}(x_i -)$, $I=1,..., n$, are asymptotically iid as *p*-variate normal with zero mean and identity covariance matrix. Tests developed by Andrews, Gnanadesikan, Mardia and others are direct functions of y_i. We observe that the *np* components of all y's put together can be considered as an asymptotically iid sample of size *np* from a univariate normal distribution with zero mean and unit variance. We tests this hypothesis using any well known test based upon independent observations for univariate normality. In particular we use univariate skewness and kurtosis tests.

Submitted: *Sankhy* .

97-22 Suman K. Mitra and C.A. Murthy. Mathematical Framework to Show the Existence of Attractor of Partitioned Iterative Function Systems.

The technique of image compression using Iterative Function System (IFS) is known as fractal image compression. An extension of IFS theory is Partitioned or local Iterative Function System (PIFS) for coding the gray level images. Several techniques of PIFS based image compression has already been proposed by many researchers. The theory of PIFS appears to be different from the theory of IFS in the sense of application domain. In the present article we have proposed a mathematical formulation for the existence of the attractor of PIFS, assuming it as a separate scheme, in the context of image compression. It has been shown that the attractor is an approximant of the given target image. The experimental results have also been presented in support of the theory. The experimental results have been obtained by using a GA based PIFS technique proposed by Mitra et al.

Submitted: *Pattern Recognition.*

97-23 G.R. Mohtashami Borzadaran and D.N. Shanbhag. Characterizations of Exponential Families via Chernoff-Type Inequalities.

Using representation theorems of Alharbi & Shanbhad (1996) and Mohtashami & Shanbhag (1997), or their variations, related to characterizations based on Chernoff-type inequalities, characterizations of exponential families are arrived at. These results subsume partial results on exponential families given earlier by Hudson (1978) and Prakasa Rao (1979).

Submitted: *London Mathematical Society.*

97-24 Jennifer Pittman and C.A. Murthy. Fitting Optimal Piecewise Linear Functions Using Genetic Algorithms.

Constructing a model for data in R^2 is a common problem in many scientific fields, including pattern recognition, computer vision, and applied mathematics. Often, little is known about the process which generated the data or its statistical properties. For example, in fitting a piecewise linear model the number of pieces as well as the knot locations may be unknown. Hence the method used to build the statistical model should have few assumptions and yet still provide a model that is optimal in some sense. Such methods can be designed through the use of genetic algorithms.

In this paper we examine the use of genetic algorithms to fit piecewise linear functions to data in R^2. The number of pieces, the location of the knots, and the underlying distribution of the data are assumed to be unknown. We discuss existing methods which attempt to solve this problem and introduce a new method which employs genetic algorithms to optimize the number and location of the linear pieces. We prove theoretically that our method provides mean-optimal functions and present the results of extensive experiments which demonstrate that the proposed method provides better results than existing spline based methods. We conclude that our method represents a valuable tool for fitting both robust and non-robust piecewise linear functions.
Submitted: *IEEE Transaction on Pattern Analysis and Machine Intelligence.*

97-25 Debasis Kundu. Estimating the Number of Sinusoids and its Performance Analysis.

Detecting the number of signals and estimating the parameters of the signals is an important problem in signal processing. Quite a number of papers

appeared in the last twenty years regarding the estimation of the parameters of the sinusoidal components but not that much of attention has been given in estimating the number of terms presents in a sinusoidal signal. Fuchs developed a criterion based on the perturbation analysis of the data auto correlation matrix to estimate the number of sinusoids, which is in some sense a subjective-based method. Recently Reddy and Biradar proposed two criteria based on AIC and MDL and developed an analytical framework for analyzing the performance of these criteria. In this paper we develop a method using the extended order modeling and singular value decomposition technique similar to that of Reddy and Biradar. We use penalty function technique but instead of using any fixed penalty function like AIC or MDL, a class of penalty functions satisfying some special properties has been used. We prove that any penalty function from that special class will give consistent estimate under the assumptions that the error random variables are independent and identically distributed with mean zero and finite variance. We also obtain the probabilities of wrong detection for any particular penalty function under somewhat weaker assumptions than that of Reddy and Biradar or Kaveh et al. It gives idea to choose the proper penalty function for any particular model. Simulations are performed to verify the usefulness of the analysis and to compare our methods with the existing ones.

Submitted: *J. Stat. Comp. & Simul.*

97-26 C.R. Rao and Shailaja Suryawanshi. Statistical Analysis of Shape through Triangulation of Landmarks: A Study of Sexual Dimorphism in Hominids.

Two objects with homologous landmarks are said to be of the same shape if the configuration of landmarks of one object can be exactly matched with that of the other by translation, rotation/reflection, and scaling. In an earlier paper, the authors proposed statistical analysis of shape by considering logarithmic differences of all possible Euclidean distances between landmarks. Tests of significance for differences in the shape of objects and methods of discrimination between population were developed using such data. In the present paper, the corresponding statistical methodology is developed by triangulation of the landmarks and considering the angles as natural measurements of shape. This method is applied to the study of sexual dimorphism in hominids.

LIST OF TECHNICAL REPORTS ISSUED BY THE CENTER FOR MULTIVARIATE ANALYSIS UNDER THE GRANT DAAH04-96-1-0082

1998

98-01 G.J. Babu, P.K. Pathak and C.R. Rao. Second Order Corrections of the Sequential Bootstrap.

Rao, Pathak and Koltchinskii (1997) have recently studied a sequential approach to resampling in which resampling is carried out sequentially one-by-one (with replacement each time) until the bootstrap sample contains m $(1-e^{-1})n$ .632n distinct observations from the original sample. They have established that the main empirical characteristics of the sequential bootstrap go through, in the sense of being within a distance of order $O(n^{-3/4})$ from those of the usual bootstrap. However, the theoretical justification of the second order corrections of the sequential bootstrap is somewhat involved. It is the main topic of this investigation. Among other things, we accomplish it by approximating our sequential scheme by a resampling scheme based on the Poisson distribution with mean $\mu=1$ and censored an X=0.

Submitted: *Annals of Statistics*.

98-02 D. Kundu. Estimating the Number of Sinusoids and its Performance Analysis.

Detecting the number of signals and estimating the parameters of the signals is an important problem in signal processing. Quite a number of papers appeared in the last twenty years regarding the estimating of the parameters of the sinusoidal components but not that much attention has been given in estimating the number of terms present in a sinusoidal signal. Fuchs developed a criterion based on the perturbation analysis of the data auto correlation matrix to estimate the number of sinusoids, which is in some sense a subjective-based method. Recently Reddy and Biradar proposed two criteria based on AIC and MDL and developed an analytical framework for analyzing the performance of these criteria. In this paper we develop a method using the extended order modeling and singular value decomposition technique similar to that of Reddy and Biradar. We use penalty function technique but instead is using any fixed penalty function line AIC or MDL, a class of penalty functions satisfying some special properties has been used. We prove that any penalty function from that special class will give consistent estimate under the assumptions that the error random variables are independent and identically distributed with mean zero and finite variance. We also obtain the probabilities of wrong detection for any particular penalty function under some what weaker assumptions that of Reddy and Biradar or Kaveh et.al.. It gives some idea to choose the proper penalty function for any particular model. Simulations are performed to

verify the usefulness of the analysis and to compare our methods with the existing ones.
Submitted: *J. Stat. Comp. & Simul.*

98-03 D. Kundu and A. Mitra. Estimating the Parameters of the Linear Compartment Model.

In this paper we consider the linear compartment model and consider the estimation procedures of the different parameters. We discuss a method to obtain the initial estimators, which can be used for any iterative procedures to obtain the least squares estimators. Four different types of confidence intervals have been discussed and they have been compared by computer simulations. We propose different methods to estimate the number of components of the linear compartment model. One data set has been used to see how the different methods work in practice.

Submitted: *J. of Stat. Planning & Inference.*

98-04 S. Bandyopadhyay, C.A. Murthy and S.K. Pal. Pattern Classification Using Genetic Algorithms: Determination of H.

A methodology based on the concept of variable string length GA (VGA) is developed for determining automatically the number of hyperplanes for modeling the class boundaries in *GA-classifier*. The genetic operators and fitness function are newly defined to take care of the variability in chromosome length. It is proved that the said method is able to arrive at the optimal number of misclassifications after sufficiently large number of iterations, and will need minimal number of hyperplanes for this purpose. Experimental results on different artificial and real life data sets demonstrate that the classifier, using the concept of variable length chromosome, can automatically evolve an appropriate value of H, and also provide performance better than those of the fixed length version. Its comparison with another approach using VGA is provided.

Submitted: *Pattern Recognition Letters.*

98-05 S. Bandyopadhyay, C.A. Murthy and S.K. Pal. Simulated Annealing Based Pattern Classification.

A method is described for finding decision boundaries, approximated by piecewise linear segments, for classifying patterns in ^{N}, N 2, using simulated annealing. It involves generation and placement of a set of hyperplanes (represented by strings) in the feature space that yields minimum misclassification. Theoretical analysis shows that as the size of the training data set approaches infinity, the boundary provided by the simulated annealing based classifier will approach the Bayes

boundary. The effectiveness of the classification methodology, along with the generalization ability of the decision boundary, is demonstrated for both artificial data and real life data sets having non-linear/overlapping class boundaries. Results are compared extensively with those of the Bayes classifier, k-NN rule and multilayer perceptron, and Genetic Algorithms, another popular evolutionary technique. Empirical verification of the theoretical claim is also provided.

Submitted: *Information Science.*

98-06 Jennifer L. Pittman. A Genetic Algorithm for Variable Knot Spline Fitting via Least Squares.

In this report we shall describe a method for fitting variable knot spline models to noisy univariate data which used a genetic algorithm to optimize over the number and location of the knots. For a fixed number of knots, the location of the knots is chosen to minimize the sum of squares error; the appropriate number of knots is determined by the adjusted GCV criterion of Luo and Wahba (1997). The objective is to find the model which minimizes *RSS/df,* where the degrees of freedom are inflated to reflect the adaptive nature of the knot search (i.e., selection of basis functions). We justify theoretically that our algorithm will converge to the variable knot model which optimizes the model fitting criterion, given that this model is contained in the search space. A modified bootstrap technique is used to obtain pointwise standard errors for models obtained by the GA method. Experimental results comparing the performance of the proposed algorithm to those obtained using the non-linear optimization technique of Schwetlick and Schütze (1995), the genetic algorithm proposed by Manela et.al. (1993), and the method of Luo and Wahba (1997). Are presented. We also discuss the extension our technique to related problems.

Submitted:

98-07 B. Uma Shankar, C.A. Murthy and S.K. Pal. A New Gray Level Based Hough Transform for Region Extraction: An Application to IRS Images.

A technique using Hough transform is described for detection of homogeneous line segments directly from (i.e., without binarization of) gray level images. A definition of "region" in terms of these line segments, with constraints on its length and variance, is provided. The algorithm is able to extract gray level regions irrespective of their shape and size. The effectiveness of the method is demonstrated on Indian Remote-sensing Satellite (IRS) images.

Submitted: *J. of Information Science.*

98-08 D. Kundu and A. Mitra. Different Methods of Estimating Sinusoidal Frequencies:

A Numerical Comparison.

In this paper we consider the problem of estimating the sinusoidal frequencies by different non iterative techniques. We briefly discuss different non iterative methods available in the literature till now to estimate the sinusoidal frequencies. We perform some simulations study and compare different methods with respect to their mean squared errors by using different models and different error distributions. It is observed that the method proposed by Quinn works very well if only one frequency is present and if the error variance is large. The modified noise space decomposition works well in other cases.

Submitted: *J. Stat. Comp. and Simul.*

98-09 C.R. Rao. Geometry of Circular Vectors and Pattern Recognition of Shape of a Boundary.

This paper deals with pattern recognition of the shape of the boundary of closed figures on the basis of a circular sequence of measurements taken on the boundary at equal intervals of a suitably chosen argument with an arbitrary starting point. A distance measure between two boundaries is defined in a such a way that it has zero value when the associated sequences of measurements coincide by shifting the starting point of one of the sequences. Such a distance measure which is invariant to the starting point of the sequence of measurements is used in identification or discrimination by the shape of the boundary of a closed figure. The mean shape of a given set of closed figures is defined and tests of significance of differences in mean shape between populations are proposed.

Submitted: *Proceedings of the National Academy of Sciences.*

98-10 M. Asadi, C.R. Rao and D.N. Shanbhag. Some Unified Characterization Results on Generalized Pareto Distributions.

In the present paper, we unify and extend various characterizations of exponential and geometric distributions, such as those based on order statistics, record values and the strong memoryless property among others, to arrive at characterization results relative to generalized Pareto distributions and their discrete versions. In the process of doing this, we also introduce here a new concept of extended neighboring order statistics, subsuming those relative to order statistics and record values in the literature without restricting ourselves to distributions that are absolutely continuous with respect to Lebesgue measure.

Submitted: *J. of Applied Probability.*

98-11 M. Bhaskara Rao and T. Ogunyemi. A Note of E-Optimal Minimal Block Designs

under Mixed Effects Model.

Mukerjee, Shah and Sinha (1992) identified the E-optimal design under mixed effects model from among the class of all minimal block designs which are connected under the fixed effects model. In this note, we show that the same design is E-optimal under the mixed effects model from among the class of all minimal block designs.

Submitted: *J. of Applied Statistical Science.*

98-12 S. Mudivarthy and M. Bhaskara Rao. Gambling as an Investment Strategy.

The focal point of the paper is an examination of whether or not gambling could be an avenue as an investment strategy for procuring a ten percent return on a given capital. Some gambling strategies are analyzed in this context. A random walk with uneven steps arises in one of the strategies.

Submitted: *American Statistician.*

98-13 S.T.S. Bukkapatnam, S.R.T. Kumara and A. Lakhtakia. Local Eigenfunctions Based Suboptimal Wavelet Packet Representation of Contaminated Chaotic Signals.

We report a suboptimal wavelet packet (WP) representation of signals emanating from a chaotic attractor contaminated by low levels of noise. Our method - geared towards choosing a suboptimal scaling function to parsimoniously represent the signal - involves extracting local eigenfunctions using artificial ensembles generated from a *pseudo-probability space*, and using the extracted local eigenfunctions to develop a suboptimal scaling function.

Submitted: *IMA J. of Applied Mathematics.*

98-14 R.D. Gupta and D. Kundu. Generalized Exponential Distributions.

The three-parameter gamma and three-parameter Weibull are the most popular distributions for analyzing any lifetime data or skewed data. Both distributions have several desirable properties and also have nice physical interpretations. In both cases, because of the scale and shape parameters, they have quite a bit of flexibility to analyze different types of lifetime data. They have increasing as well as decreasing hazard rate depending on the shape parameter. Unfortunately both distributions have certain drawbacks also. In this paper we consider a three-parameter distribution which is a particular case of the exponentiated Weibull distribution originally proposed by Mudholkar et al. (1995) when the location parameter is not present. We study different properties of this model and observe that this family has some interesting features which are quite similar to that of the gamma family and the

Weibull family and it has certain distinct properties also. We believe this model can be used as an alternative to the gamma model or the Weibull model in many situations. We provide one data set where three-parameter generalized exponential distribution fits better than three-parameter Weibull distribution or three-parameter gamma distribution.

Submitted: *Australian J. of Statistics.*

98-15 D. Kundu. Estimating Direction of Arrival of Signals: Some Asymptotic Results.

In signal processing high resolution signal parameter estimation is a significant problem. In particular the estimation of the direction of arrival of narrow band signals emitted by multiple sources received wide applications recently in signal processing literature. It is well known that among several methods, the centro symmetric TLS-ESPRIT and MUSIC work very well. In this paper we propose two new eigendecomposition methods, which work better than centro symmetric TLS-ESPRIT and MUSIC in terms of lower mean squared errors at least for small SNR and computational costs of both the methods are less than that of MUSIC or centro symmetric TLS-ESPRIT. Some asymptotic results have been obtained.

Submitted: *Sankhy , A.*

98-16 D. Kundu and A. Mitra. Fitting a Sum of Exponentials to Equispaced Data.

In this paper we consider fitting a sum of exponentials to equispaced data. We propose a new non iterative method to estimate the parameters, which can be used as an initial guess in any standard minimization algorithm. We use two real life data sets to see how the new estimate behaves as an initial guess for different algorithms available in the literature. It is well known (Wu; 1981) that for the exponential model any estimate is inconsistent in the usual sense, we redefine the consistency and the asymptotic normality of the least squares estimates under this new assumption.

Submitted: *Sankhy , A.*

98-17 R.D. Gupta and D. Kundu. Hybrid Censoring Schemes with Exponential Failure Distribution.

The mixture of Type I and Type II censoring schemes, called the hybrid censoring, is quite important in life-testing experiments. Epstein (1954, 1960) introduced this testing scheme and proposed a two-sided confidence interval to estimate the mean lifetime, , when the underlying lifetime distribution is assumed to be exponential. There are some two-sided confidence intervals and credible intervals proposed by Fairbanks et al. (1982) and Draper and Guttman (1987) respectively. In this paper we obtain the exact two-sided confidence interval of following the approach of

Chen and Bhattacharya (1988). We also obtain the asymptotic confidence intervals in the Hybrid censoring case. It is important to observe that the results for Type I and Type II censoring schemes can be obtained as particular cases of the Hybrid censoring scheme. We analyze one data set and compare different methods by Monte Carlo simulations.

Submitted: *Communications in Stat., A.*

98-18 G.J. Babu, P.K. Pathak and C.R. Rao. Consistency and Accuracy of the Sequential Bootstrap.

The object of this paper is to present a brief account of the sequential bootstrap from a survey sampling point of view. This sequential resampling scheme entails resampling from the observed sample sequentially (with replacement) until a preassigned number of distinct original observations appear. This approach stems from the observation made by Efron (1983) that the usual bootstrap samples are supported on approximately .632n of the original data points. We outline a number of approaches that can be employed to study the theoretical as well as the empirical properties of the sequential bootstrap. Our investigation shows that there is a great potential for sequential bootstrap in applications often encountered in practice.

Submitted: *Annals of Statistics.*

98-19 H. Toutenburg, A. Fieger and V.K. Srivastava. Weighted Modified First Order Regression Procedures for Estimation in Linear Models with Missing - Observations.

This paper considers the estimation of coefficients in a linear regression model with missing observations in the independent variables and introduces a modification of the standard first order regression method for imputation of missing values. The modification provides stochastic values for imputation and, as an extension, makes use of the principle of weighted mixed regression. The proposed procedures are compared with two popular procedures - one which utilizes only the complete observations and the other which employs the standard first order regression imputation method for missing values.
A simulation experiment to evaluate the gain in efficiency and to examine interesting issues like the impact of varying degree of multicollinearity in explanatory variables is proceeded. Some work on the case of discrete regressor variables is in progress and will be reported in a future article to follow.

Submitted: *Statistical Papers.*

98-20 B. Schaffrin and H. Toutenburg. The Impact of Missing Values on the Reliability Measures in a Linear Model.

Reliability measures in linear models are used in geodetic science and elsewhere to quantify the potential to detect outliers and to suppress their impact on the regression estimates. Here we shall study the effect of missing values on these reliability measures with the idea that, under a proper design, they should not change drastically when such a situation occurs.

Submitted: *Computational Statistics and Data Analysis.*

98-21 D. Kundu and S. Basu. Analysis of Incomplete Data in Presence of Competing Risks.

In medical studies or in reliability analysis an investigator is often interested with the assessment of a specific risk in presence of other risk factors. In the Statistical literature it is known as the analysis of competing risks model. The competing risks model assumes that the data consists of a failure time and an indicator denoting the cause of failure. Several studies have been carried out under this assumption for parametric and non parametric set up. Unfortunately in many situations, the causes of failure are not observed, even if the failure time is observed. Miyawaka (1984) obtained some of the results under the assumption that the failure time distribution is exponential. He obtained the maximum likelihood estimators and the minimum variance estimators of the unknown parameters. We provide the approximate and asymptotic properties of these estimators. Using the approximate and the asymptotic distributions we obtain confidence bounds of the parameters and also propose two different bootstrap confidence bounds. We consider the case when the failure distribution may not be exponential and use one data set to see how different methods work in real life situations.

Submitted: *J. Statistical Planning and Inference.*

98-25 C. Radhakrishna Rao. Statistical Proofs of Some Matrix Inequalities.

Matrix algebra is extensively used in the study of linear models and multivariate analysis (see for instance Rao (1973) and Rao and Rao (1998)). During recent years, there have been a number of papers where statistical results are used to prove some matrix theorems, especially matrix inequalities (Mitra (1996), Mitra and Puntanen (1991), Dey, Hande and Tikku (1994), Kagan and Landsman (1996), Kagan and Smith (1998), etc.). In this paper, a number of matrix results are proved using some properties of Fisher information and covariance matrices. A unified approach is provided through the use of Schur complements. It may be noted that the statistical results used are derivable without using matrix theory.

Submitted: *Linear Algebra and its Applications.*

LIST OF TECHNICAL REPORTS ISSUED BY THE CENTER FOR MULTIVARIATE ANALYSIS UNDER THE GRANT DAAH04-96-1-0082

1999

99-01 Z.J. Liu and C.R. Rao. MU-estimation and Smoothing.

In the traditional M-estimation theory developed by Huber (1964), the parameter under estimation is the value of which minimizes the expectation of what is called a discrepancy measure (DISM) (x,) which is a function of and the underlying random variable X. Such a setting does not cover the estimation of parameters such as multivariate median defined by Oja (1983) and Liu (1990), as the value of which minimizes the expectation of a DISM of the type (X_1 ,...,X_m,) where X_1 ,...,X_m are independent copies of the underlying random variable X. Arcones et al (1994) studied the estimation of such parameters. We call the associated M-estimation MU-estimation (or μ-estimation for convenience). When a DISM is not a differentiable function of , some complexities arise in studying the properties of estimations as well as in their computation. In such a case, we introduce a new method of smoothing the DISM with a kernel function and using it in estimation. It is seen that smoothing allows us to develop an elegant approach to the study of asymptotic properties and computation of estimations.

Submitted: *J. Multivariate Analysis.*

99-02 C.R. Rao. Pre and Post Least Squares: The Emergence of Robust Estimation.

The paper traces the history of estimation of unknown parameters when measurements are subject to error from the time of Ptolemy to Gauss and Laplace, the inventors of the method of least squares estimation (LSE). The modern theory of least squares started with papers by Markov and Aitken on estimation of parameters in a linear regression model and Fisher on tests of significance of regression coefficients. Some refinements in the general setup of a linear model and the associated statistical methodology were made by Bose and Rao.

The LSE's have some nice properties; however, they are found to be sensitive to outliers and contamination in data. To overcome this defect, robust methods are introduced by minimizing the sum of nonnegative functions of residuals which increase at a slower rate than the squares. Some attempts to build a unified theory of robust estimation are discussed.

Submitted: *J. of Stat. Planning and Inference..*

99-03 G. Jogesh Babu, P.K. Pathak, and C.R. Rao. Second Order Corrections of the Poisson Bootstrap.

Rao, Pathak and Koltchinskii (1997) have recently studied a sequential approach to resampling in which resampling is carried out sequentially one-by-one (with replacement each time) until the bootstrap sample contains m $(1-e^{-1}\})n$ .632n distinct observations from the original sample. In our previous work, we have established that the main empirical characteristics of the sequential bootstrap go through, in the sense of being within a distance of order $O(n^{-3/4})$ from those of the usual bootstrap. However, the theoretical justification of the second order correctness of the sequential bootstrap is somewhat difficult. It is the main topic of this investigation. Among other things, we accomplish it by approximating our sequential scheme by a resampling scheme based on the Poisson distribution with mean μ and censored at X=0.

Submitted: *Annals of Stat.*.

99-05 R.D. Gupta and D. Kundu. Feasibility of the Method of Moment Estimators.

In this paper we consider the problem of feasibility of the method of moment estimators. We consider three different distributions, namely exponential, uniform and generalized Pareto. In some cases, we obtain the explicit probability of the feasible method pf moment estimator and other cases we estimate it through simulations. Since the method of moment estimators are always strongly consistent, it is expected that the feasibility probability should also increase to one. But some of the counter intuitive results are obtained. In some cases it is observed that the probability may even converge to zero also.

Submitted: *J. of Mathematical and Statistical Sciences.*

99-06 D. Kundu and R.D. Gupta. A Note on the Asymptotic Behavior of the LSE'S if the Parameters for Superimposed Exponential Signals in Presence of Stationary Noise.

Superimposed exponential signals play an important role in Statistical Signal Processing and Time series analysis. In this note, the asymptotic behavior of the least squares estimators of the parameters are obtained in presence of stationary noise for the undamped exponential model. It is well known that this model does not satisfy the sufficient conditions of Jennrich (1969), Wu (1981) or Kundu (1991) for the least squares estimators to be consistent even when the errors are independent and identically distributed random variables with mean zero and finite variance. This paper extends some of the earlier works of Hannan (1971, 1973), Walker (1971), Bai et al. (1991), Rao and Zhao (1993), Kundu (1995) and Kundu and Mitra (1995, 1998) in different ways. Some numerical experiments are performed to observe the small sample behavior of the least squares estimators.

Submitted: *American J. of Mathematical and Management Sciences.*

99-07 S. Nandi and D. Kundu. On Asymptotic Properties of the Least Squares Estimators in a Stationary Random Field.

A particular two dimensional model in a stationary random field, which has a wide applications in statistical signal processing and in texture classifications, is considered. We prove the consistency and also obtain the asymptotic distributions of the least squares estimators of the different model parameters. It is observed that the asymptotic distribution of the least squares estimators are multivariate normal. Some numerical experiments are performed ti see how the asymptotic results work for finite samples. We propose some open problems at the end.

Submitted: *Multidimensional Systems and Signal Processing..*

99-08 S.K. Mitra, C.A. Murthy and M.K. Kundu. A Technique for Image Magnification Using Partitioned Iterative Function System.

A new technique for image magnification using the theory of fractals is proposed. The technique is designed assuming self transformability property of images. In particular, the magnification task is performed using the fractal code of the image instead of the original one resulting in a reduction in memory requirement. To generate the fractal codes, Genetic Algorithm with elitist model is used which greatly decreases the search for finding self similarities in the given image. The article presents both theory and implementation of the proposed method. A simple distortion measure scheme and a similarity criterion based on just noticeable difference have also been proposed to judge the image quality of the magnified image. Comparison with one of the most popular magnification techniques, the nearest neighbor technique, is made.

Submitted: *Pattern Recognition..*

99-09 Suman K. Mitra and C.A. Murthy. Mathematical Framework to Show the Existence of Attractor of Partitioned Iterative Function Systems.

The technique of image compression Iterative Function System (IFS) is known as fractal image compression. An extension of IFS theory is Partitioned or local Iterative Function System (PIFS) for coding the gray level images. Several techniques of PIFS based image compression have already been proposed by many researchers. The theory of PIFS appears to be different from the theory of IFS in the sense of application domain. The present article discusses some basic differences between IFS and PIFS and provides a separate mathematical formulation for the existence of attractor of partitioned IFS. In particular, it has been shown that the attractor exists and it is an approximation of the given target image. The experimental results have also been presented in support of the theory. The experimental results have been obtained by using a GA based PIFS technique proposed by Mitra et al [1].

Submitted: *Pattern Recognition.*

99-10 V. Seshadri and D.N. Shanbhag. An Extended Laha-Lukacs Characterization Result

Revisited.

Recently Fosam and Shanbhag (1997) gave an extended version of the Laha-Lukacs characterization result based on a regression property, subsuming the Letac-Mora characterizations of the natural exponential families of distributions with variances as cubic functions of means. In the present note, we provide a new approach based on functional equations to arrive at the Fosam-Shanbhag result.

Submitted: *J. of Theoretical Probability.*

99-11 Majid Asadi and D.N. Shanbhag. Hazard Measure and Mean Residual Life Ordering: A Unified Approach.

The hazard rate ordering is applied frequently in reliability to compare two probability distributions on R_+ such that they are both absolutely continuous (w.r.t. Lebesgue measure) or both purely discrete (concentrated on the set of non-negative integers) via their hazard rates. Kotz and Shanbhag (1980) extended the concept of hazard rate introducing new concept of hazard measure, applicable to any arbitrary distribution on the real line; in particular, this concept avoids the restriction that the distribution be absolutely continuous or purely discrete. These authors have also extended the concept of mean residual life function and have given related representations for distributions. In this paper, we introduce the concepts of hazard measure ordering and mean residual life ordering to compare two arbitrary probability distributions and study their basic properties.

Submitted: *Annals of Applied Probability.*

99-12 Debasis Kundu. On the Determination of the Number of Signals and its Performance Analysis in Presence of White Noise.

In signal processing, high resolution signal parameter estimation is a significant problem. In particular the estimation of the direction of the narrow band signals emitted by multiple sources received wide applications recently in signal processing literature. Quite a number of papers appeared in the last twenty five years regarding the estimation of the parameters of the direction of arrival of signals, but not that much attention has been given in estimating the number of signals. In this paper we develop a method using penalty function technique. But instead of using any fixed penalty function like AIC or MDL, a class of penalty functions satisfying some special properties have been used. We proved that any penalty function from that particular class will produce consistent estimates under the assumptions that the error random variables are independent and identically distributed with mean zero and finite variance. We also obtain the probabilities of wrong detection for any particular penalty function and estimate it using the matrix perturbation technique. It gives some idea to choose the proper penalty function for any particular model. Simulations are performed to verify the usefulness of the analysis and to compare our method with the existing ones.

Submitted: *Signal Processing.*

99-13 Y. Wu, K.W. Tam, F. Li and M.M. Zen. A Note on estimating the Number of Super Imposed Exponential Signals by the Cross-Validation Approach.

In this paper, a procedure based on the delete-1 cross-validation is given for estimating the number of super imposed exponential signals, its limiting behavior is explored and it is shown that the probability of overestimating the true number of signals is larger than a positive constant for sample size large enough. Also a general procedure based on the cross-validation is presented when the deletion precedes according to a collection of subsets of indices. The result is similar to the delete-1 cross-validations if the number of deletion is fixed. The simulation results are provided for the performance of the procedure when the collections of subsets of indices are chosen as those suggested by J. Shao in a linear model selection problem.

Submitted: *IEEE Transactions on Signal Processing.*

99-14 D. Kundu and A. Mitra. Detecting the Number of Signals for an Undamped Exponential Model Using Cross Validation Approach.

Detecting the number of signals and estimating the parameters of the signals are important problems in Statistical Signal Processing. Quite a number of papers appeared in the last twenty years in estimating the parameters of an exponential signal quite efficiently but not that much of attention has been paid in estimating the number of signals of an exponential signal model. Recently it is observed that different Information Theoretic Criteria can be used to estimate the number of signals in this situation. But it is also observed that the choice of the penalty function is very important particularly for small sample sizes. In this paper we suggest to use the Cross Validation technique on estimating the number of signals and give its practical implementation procedures. Numerical experiments reveal that the new procedure performs quite comparable to the best performed Information Theoretic Criteria at least for small sample sizes and it has certain desirable properties also.

Submitted: *J. of Indian Institute Science.*

99-15 S. Nandi and D. Kundu. Least Squares Estimators in a Stationary Random Field.

A particular two dimensional model in a stationary random field, which has wide applications in statistical signal processing and in texture classifications, is considered. We prove the consistency and also obtain the asymptotic distributions of the least squares estimators of the different model parameters. It is observed that the asymptotic distribution of the least squares estimators are multivariate normal. Some numerical experiments are performed to see how the asymptotic results work for finite samples. We propose some open problems at the end.

Submitted: *Signal Processing.*

99-16 D. Kundu and S. Basu. Analysis of Incomplete Data in Presence of Competing Risks.

In medical studies or in reliability analysis an investigator is often interested in the assessment of a specific risk in presence of other risk factors. In the Statistical literature it is known as the analysis of competing risks model. The cometing risks model assumes that the data consists of a failure time and an indicator denoting the cause of failure. Several studies have been carried out under this assumption for parametric and non parametric set up. Unfortunately in many situations, the causes of failure are not observed, even if the failure times are observed Miyawaka (1984) obtained some of the results under the assumption that the failure time distribution is exponential. He obtained the maximum likelihood estimators and the minimum variance unbiased estimators of the unknown parameters. We provide the approximate and asymptotic properties of these estimators. Using the approximate and the asymptotic distributions we compute confidence intervals of the parameters and compare them with the two different bootstrap confidence bounds. We also consider the case when the failure distributions are Weibull. One data set is used to see how different methods work in real life situations.

Submitted: *JSPI*

99-17 D. Kundu and A. Mitra. Estimating the Number of Signals of the Damped Exponential Models.

In this paper, we consider the problem of estimating the number of signals of the damped exponential models. We use different information theoretic criteria to detect the number of signals and compare their small sample performances by Monte Carlo simulations study.

Submitted: *Computational Statsitics and Data Analysis*

Appendix E: Participation in Conferences During the Period 1996-1999 by C.R. Rao

1996

83rd Session of the Indian Science Congress Association, Patiala, Jan.3, 1996.
Invited talk: Cross examination of data: the art of making figures speak.

Symposium on Estimating Functions, University of Georgia, Athens, USA, March 21-23, 1996.
Invited talk.- Inference on regression models using the difference of two convex functions as the discrepancy measure (21 March).

US Army Conference on Applied Statistics, Monterey, California, October 23-25, 1996.
Keynote Address: Pre and Post Least Squares: The Emergence of Robust Estimation (October 23).

Conference on Statistical Research in the 21st Century, An IISA (International Indian Statistical Association) in honor of C.R. Rao, Montreal, Canada, Nov 29-30, 1996.
Keynote Address: Pre and Post Least Squares: The Emergence of Robust Estimation (Nov 29).

1997

International Symposium on Contemporary Multivariate Analysis and its Applications, Hong Kong, May 19-22, 1997.
Keynote Address: An alternative to correspondence analysis using Hellinger distance, May 22, 1997.

The Third Conference of the Forum for Interdisciplinary Mathematics: International Conference on Combinatorics, Information Theory and Statistics, Portland, Maine, USA, July 18-20, 1997.
Keynote Address: Pre and post least squares: The emergence of robust estimation, 19 July 1997.

Sixth International Workshop on Matrix Methods in Statistics. Istanbul, 16-17 August 1997.
Invited talk: Statistical solutions to some matrix problems, 16 Aug, 1997.

51st Session of the International Statistical Institute, Istanbul, 18-26 August 1997.
Invited talk: A cross disciplinary approach to teaching of statistics, in the session on data centered versus mathematics centered training in statistics, 20 Aug, 1997.

1998

85th Session of the Indian Science Congress, Hyderabad, Jan 3-10.
Invited talk: Pre and post least squares: The emergence of robust estimation, Jan 5.

Workshop on Computer and Communication Technologies in Education, Jan 19-20, Indian Statistical Institute, Calcutta.
Invited talk: On the use of computers in education, Jan 19.

Conference on Statistical Methods for Quality and Productivity Improvement: Current Trends, Feb 5-6, Indian Statistical Institute, New Delhi.
Keynote Address: Statistics for total quality management, Feb 5.

A Symposium on Statistics in Celebration of the First Anniversary of the Department of Statistics, University of Pittsburgh, March 27-28, 1998.
Inaugural lecture: Statistics: A technology for the Millennium, March 27.

1998 Lukacs Symposium: Statistics for the 21st Century, April 24-26, 1998.
Inaugural lecture: A review of canonical coordinates for reduction of multivariate data and an alternative to correspondence analysis, April 24.

Rutgers Bootstrap Conference, Rutgers University, USA, 14-16 May 1998.
Invited talk: Second order corrections of the sequential bootstrap, May 16.

Seventh International Workshop on Matrices & Statistics, Fort Lauderdale, Florida, Dec 11-14, 1998.
Keynote lecture: Statistical solutions to some matrix problems, Dec. 12.

6th Escola de Modelos de Regressao (The Sixth School of Regression Model), Feb 8-10, 1999, Brasilia (Brazil).
Inaugural Address: Statistics: A Technology for the Millenium, 8 Feb.

Frontiers of Environmental and Ecological Statistics for the 21st Century. Ninth Lukacs Symposium, Bowling Green State University, April 23-25, 1999.
Panel Discussion: Issues and Approaches for Combining Practicality with Cross-Disciplinary Scholarship into the New Millennium with Floor Discussion.
Contribution: Statistics and Data Mining.